This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT/CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau

PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : A01N 43/22, A61K 31/71 C07H 17/08, C12N 1/21 C12P 19/62	A1	`	1) International Publication Number: 3) International Publication Date:	WO 93/13663 22 July 1993 (22.07.93)
(21) International Application Number: PCT/US (22) International Filing Date: 17 January 1992			(81) Designated States: AU, CA, JP, 1 BE, CH, DE, DK, ES, FR, G SE).	
(71) Applicant: ABBOTT LABORATORIES [US/US D-377/AP6D-2, One Abbott Park Road, Abbot 60064-3500 (US).	ij; CH/ tt Park,	AD IL	Published With international search report	ı.
(72) Inventors: KATZ, Leonard; 844 North Avenue gan, IL 60085 (US). DONADIO, Stefano; 110 Road, Libertyville, IL 60048 (US). MCALPIN B.; 211 West Rockland Road, Libertyville, (US).	Brook! E, Jam	hill ies,		
(74) Agents: GORMAN, Edward, Hoover, Jr. et al.; A boratories, CHAD D-377/AP6D-2, One Ab Road, Abbott Park, IL 60064-3500 (US).				

(54) Title: METHOD OF DIRECTING BIOSYNTHESIS OF SPECIFIC POLYKETIDES

(57) Abstract

A method to produce novel polyketide structures by designing and introducing specified changes in the DNA governing the synthesis of the polyketide is disclosed. The biosynthesis of specific polyketide analogs is accomplished by genetic manipulation of a polyketide-producing microorganism by isolating a polyketide biosynthetic gene-containing DNA sequence, identifying enzymatic activities associated within the DNA sequence, introducing one or more specified changes into the DNA sequence which codes for one of the enzymatic activities which results in an altered DNA sequence, introducing the altered DNA sequence into the polyketide-producing microorganism to replace the original sequence, growing a culture of the altered microorganism under conditions suitable for the formation of the specific polyketide analog, and isolating the specific polyketide analog from the culture. The method is most useful when the segment of the chromosome modified is involved in an enzymatic activity associated with polyketide biosynthesis, particularly for manipulating polyketide synthase genes from Saccharapolyspora or Streptomyces.





FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Amtria	FR	France	MR	Mauritania
		GA	Gabon	MW	Malawi
UA	Australia			NL	Netherlands
BB	Barbados	GB	United Kingdom		
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	KU	Hungary	PL	Poland
BJ	Benin	ΙE	Ireland	PT	Portugal:
BR	Brazil	iT	Italy	RO	Romania .
CA	Canada	JP	Japan-	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic	SD	Sudan
CC	Congo		of Korca	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SK	Slovak Republic
CI	C'ôte d'Ivoire	KZ	Kazakhstan	SN	Senegal
CM	Canuroon	IJ	Liechtenstein	SU	Soviet Union
CS	Czechoslovakia	LK	Sri tanka	TD	Chad
CZ	Czech Republic	1.0	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	UA	Ukraine
DK	Denmark	MG	Madagascar	US	United States of America
ES	Spain	ML.	Mali	VN	Viet Nam
FI	Finland	MN	Mangolia		

PCT/US92/00427

5

15

20

25

30

35

METHOD OF DIRECTING BIOSYNTHESIS OF SPECIFIC POLYKETIDES

Field of the Invention.

The present invention relates to a method for directing the biosynthesis of specific polyketide analogs by genetic manipulation. In particuar, polyketide biosynthetic genes are manipulated to produce precise, novel polyketides of predicted structure.

10 Background of the Invention

Polyketides are a large class of natural products that includes many important antibiotics and immunosuppressants such as erythromycins, tetracyclines, and rapamycins. Their synthesis proceeds by an ordered condensation of acyl esters to generate carbon chains of varying length and substitution pattern that are later converted to mature polyketides. This process has long been recognized as resembling fatty acid biosynthesis, but with important differences. Unlike a fatty acid synthase, a typical polyketide synthase is programmed to make many choices during carbon chain assembly: For example, the choice of "starter" and "extender" units, which are often selected from acetate, propionate or butyrate residues in a defined sequence. The choice of using a full cycle of reductiondehydration-reduction after some condensation steps, omitting it completely, or using one of two incomplete cycles (reduction alone or reduction followed by dehydration), which determines the pattern of keto or hydroxyl groups and the degree of saturation at different points in the chain is additionally programed. Finally the choice of stereochemistry for the substituents at many of the carbon atoms is programmed by the polyketide synthase.

Because of the commercial significance of *Streptomyces*, a great amount of effort has been expended in the study of *Streptomyces* genetics. Consequently much is known about *Streptomyces* and several cloning vectors exist for transformations of the organism.

Although many polyketides have been identified, there remains the need to obtain novel polyketide structures with enhanced properties. Current methods of obtaining such molecules include screening of natural isolates and chemical modification of existing polyketides, both of which are costly and time consuming. Current screening methods are based on gross properties of the molecule, i.e. antibacterial, antifungal

activity, etc., and both a priori knowledge of the structure of the molecules obtained or predetermination of enhanced properties are virtually impossible. Chemical modification of preexisting structures has been successfully employed, but it still suffers from practical limitations to the type of compounds obtainable, largely connected to the poor yield of multistep syntheses and available chemistry to effect modifications. The following modifications are extremely difficult or inefficient at the present time: change of the stereochemistry of the side chains in the completed polyketide; change of the length of the polyketide by removal or addition of carbon units from the interior of the acyl chain; and dehydroxylation at unique positions in the acyl chain. Accordingly, there exists the need to obtain molecules wherein such changes can be specified and performed and would represent an improvement in the technology to produce altered polyketide molecules with predicted structure.

15

20

30

35

10

Summary of the Invention

The present invention provides a method to produce novel structures from designing and introducing specified changes in the DNA governing the synthesis of the polyketide. According to the method of the present invention, the biosynthesis of specific polyketide analogs is accomplished by genetic manipulation of a polyketide-producing microorganism comprising the steps of:

- (1) isolating a polyketide biosynthetic gene-containing DNA 2.5 sequence;
 - (2) identifying enzymatic activities associated within said DNA sequence;
 - (3) introducing one or more specified changes into said DNA sequence which codes for one of said enzymatic activities which results in an altered DNA sequence;
 - (4) introducing said altered DNA sequence into the polyketideproducing microorganism to replace the original sequence;
 - (5) growing a culture of the altered microorganism under conditions suitable for the formation of the specific polyketide analog; and
 - (6) isolating said specific polyketide analog from the culture.

The present method is most useful when the segment of the chromosome modified is involved in an enzymatic activity associated with polyketide biosynthesis. The present invention is especially useful





10

1.2

in manipulating polyketide biosynthetic genes from *Streptomyces*, an organism which provides over one-half of the clinically useful antibiotics.

Brief Description of the Drawings

FIG. 1 illustrates the organization of gene encoding polyketide synthase and designated eryA as follows: (a) Map coordinates of the DNA; (b) DOTPLOT of the output of COMPARE (window = 50, stringency = 32) program (Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin, Biotechnology Center, 1710 University Avenue, Madison, Wisconsin 53705 of eryA segment (x-axis) vs. subsegment of eryA comprises between 23 - 27.5 sequence coordinates (y-axis) [see Fig. 2]; (c) Open reading frame organization of eryA and enzymatic activities encoded. PT = propionyltransferase; ACP = acyl carrier protein; KS= β -ketoacyl ACP synthase; RmT = (2R) methylmalonyl CoA transferase; KR = β -ketoredu mase; SmT = (2S) methylmalonyl CoA transferase; DH = dehydratase; ER = enoylreductase; TE = thioesterase; and (d) Schematic diagram showing the extent of each of the six modules in eryA.

FIG. 2. illustrates the nucleotide sequence of eryA with

20 corresponding translation of the three open reading frames. Standard one
letter codes for the amino acids appear beneath their respective nucleic
acid codons. The standard one letter codes for the amino acid sequences
are as follows:

25	Α	-alanine
•	R	-arginine
	N	-asparagine
	D	-aspartic acid
	С	-cysteine
30	Q	-glutamine
	E	-glutamic acid
	G	-glycine
-	H	-histidine
	I	-isoleucine
3 5	L	-leucine
	K	-lysine
	M	-methionine (start)
	F	-phenylalanine

PCT/US92/00427

5

10

15

P	-proline
S	-serine
T	-threonine
W	-tryptophan
Y	-tyrosine
v	-valine

FIG. 3. is a schematic representation of Type I, Type II and Type III changes in eryA and structures of corresponding novel polyketides produced. $\Delta 69$ (Type I) and $\Delta 33$ (Type II) represent in-frame deletions of the base pairs in the DNA segments corresponding to the KR of module 2 and the β -ketoacyl ACP synthase of module 2, respectively. Insertion of a complete copy of module 4 within module 1 is also shown. Production of 11-epifluoro-15-norerythromycin in strain that carries $\Delta 33$ occurs when substrate analog (2S,3S,4S,5S)2,4-dimethyl-3-fluoro-5-hydroxyhexanoic acid-ethyl thioester is fed.

FIG. 4 illustrates the restriction site coordinates of cosmid pR1 5' to the sequence of eryA (Fig 2).

20

25

30

35

Detailed Description of the Invention

For the purposes of the present invention as disclosed and claimed herein, the following terms are defined.

The term "polyketide" as used refers to a large and diverse class of natural products, including antibiotics, pigments, and immunosuppressants. Antibiotics include, but are not limited to anthracyclines, tetracyclines, polyethers, ansamycins, macrolides of different types (polyenes and avermectins as well as classical macrolides such as erythromycins).

The term "polyketide-producing microorganism" as used herein includes any Actinomycetales which can produced a polyketide. Examples of Actinomycetes that produce polyketides include but are not limited to Micromonospora rosaria, Micromonospora megalomicea, Sacharapolyspora erythraea, Streptomyces antibioticus, Streptomyces albireticuli, Streptomyces ambofasciens, Streptomyces avermitilis, Streptomyces fradiae, Streptomyces hygroscopicus, Streptomyces tsukubaensis, Streptomyces griseus, Streptomyces mycarofasciens, Streptomyces platensis, Streptomyces venezuelae, Streptomyces

10

15

20

25

35

violaceoniger, and various Actinomadura, Dactylosporangium and Nocardia strains that produce polyether type of polyketides.

The term "polyketide synthase" as used herein refers to the complex of enzymatic activities responsible for the biosynthesis of polyketides which include but are not limited to β -ketoreductase, dehydratase, acyl carrier protein, enoylreductase, β -ketoacyl ACP synthase, and acyltransferase.

The term "extender" as used herein refers to a coenzyme A thioester of a dicarboxylate which is incorporated into a polyketide by a polyketide synthase.

The term "starter" as used herein refers to a coenzyme A thioester of a carboxylic acid which is used by the polyketide synthase as the first building block of the polyketide.

The term "eryA" as used herein refers to the genes involved in the formation of the polyketide moiety of erythromycin.

The term "condensation" as used herein refers to the addition of an extender unit out to the nascent polyketide chain and requires the action of β -ketoacyl ACP synthase, acyltransferase, and acyl carrier protein.

The term " β -carbonyl processing" as used herein refers to changes effecting the carbonyl group of the growing polyketide via β -ketoreductase, dehydratase, and enoylreductase.

The term "module" as used herein refers to the genetic element encoding one condensation step, as defined above, and one β -carbonyl processing step, as defined herein.

The term "Type I change" as used herein refers to changes in DNA sequence which will result in the production of polyketide rings of length identical to that of 6-deoxyerythronolide A, but with altered functional groups at specific ring positions.

The term "Type II change" as used herein refers to alterations

which will result in the production of macrolide rings only when fed exogenously with substrate analogs, e.g. thioesters of appropriate acyl compounds of various length. Thus Type II mutants are erythromycin non-producing (Ery⁻) mutants. The structure of the resulting macrolides will depend on the substrate employed.

The term "Type III change" as used herein refers to alterations which will result in the biosynthesis of macrolide rings of length reduced (deletion) or increased (insertion) by two carbon units, or macrolide rings altered in specific portions of the chain (replacement).

10

15

25

In its broadest sense, the present invention entails a general procedure for producing novel polyketide structures in vivo by selectively altering the genetic information of the organism that naturally produces a related polyketide. A set of examples described herein are a series of novel polyketides that make use of the genetic information for the biosynthesis of the polyketide portion of the macrolide antibiotic erythromycin. The organization of the segment of the Saccharapolyspora erythraea chromosome, designated eryA, and the corresponding polypeptides which it encodes that determine the biosynthesis of the polyketide segment of erythromycin, are shown in FIG. 1. It is seen that ery A is organized in modules, as shown, and that each module takes care of one condensation step, through the action of the β-ketoacyl ACP synthase specified within, wherein an extender unit, methylmalonyl CoA, is added first to the starter unit, propionyl CoA, and then to the successively growing acyl chain. The precise succession of elongation steps is dictated by the genetic order of the six modules: module 1 determines the first condensation; module 2, the second; module 3, the third, and so on until the sixth condensation step has occurred. Furthermore, the processing of the growing chain after each condensation is also determined by the information within each module. Thus \$-20 ketoreduction of the β-carbonyl takes place after each step except for step 3, as determined by the presence of a functional β -ketoreductase in all modules except module 3, whereas dehydration and enoyireduction only take place after the fourth extender unit is added to the growing acyl chain, as determined by the presence of dehydratase and enoylreductase in module 4. Furthermore, the choice of the correct enantiomer (2R or 2S) of methylmalonyl-CoA as the extender unit employed at each condensation is specified by the acyltransferase function determined by each module (FIG. 1C).

In the present invention, novel polyketide molecules of desired 30 structure are produced by the introduction of specific genetic alterations of the eryA sequence into the Sac. erythraea chromosome. The complete nucleotide sequence of the eryA segment of the Sac. erythraea chromosome and the sequence of the corresponding polypeptides are shown in FIG. 2. Three types of alterations to the eryA DNA sequence are 35 described: (i) those inactivating a single function in a module which does not arrest acyl chain growth (β -ketoreductase, dehydratase, or enoylreductase); (ii) those inactivating a single function in a module



10

15

20

25

30

35

which does arrest chain growth (β-ketoacyl ACP synthase, acyltransferase or acyl carrier protein); and (iii) those affecting an entire module (deletion, insertion, or replacement). The novel polyketides produced by strains carrying these types of mutations can be classified accordingly. Type I changes will result in the production of polyketide rings of length identical to that of 6-deoxyerythronolide A, but with altered functional groups at specific ring positions. Strains carrying type II alterations will result in the production of macrolide rings only when fed exogenously with substrate analogs, e.g.thioesters of appropriate acyl compounds of various length. Thus Type II mutants are erythromycin non-producing (Ery-) mutants. The structure of the resulting macrolides will depend on the substrate employed. Type III changes will result in the biosynthesis of macrolide rings of length reduced (deletion) or increased (insertion) by two carbon units, or macrolide rings altered in specific portions of the chain (replacement). A schematic representation of some examples of Type I, Type II and Type III alterations in eryA and the corresponding novel polyketides produced in hosts that carry such alterations is shown in FIG. 3.

In the examples described herein, specific mutations in the eryA region of the Sac. erythraea chromosome are introduced by a simple twostep approach: 1) introduction of a specified change in a cloned DNA segment; 2) exchange of the wild type allele with the mutated one. Step 1 requires standard recombinant DNA manipulations employing E. coli as the host. Step 2 requires one or more plasmids out of the several E. coli-Sac. erythraea shuttle vectors available and a simple screening procedure for the presence of the colony carrying the altered gene. Two methods are used to introduce the altered allele into the chromosome to replace the wild type allele. The first employs gene replacement, described in Examples 7, 11, 15, 19 and 24, wherein the gene to be altered, along with adjacent upstream and downstream DNA, is mutated and cloned into a Sac. erythraea non-replicating vector. The plasmid carrying the altered allele is then introduced into the host strain by transformation of protoplasts employing selection for a plasmid marker. Since the plasmid does not replicate, regenerated cells that carry the marker have undergone a single homologous recombination between one of the two segments flanking the mutation on the plasmid and its homologous counterpart in the chromosome. Some of the colonies that have subsequently lost the marker will have undergone a second recombination between the other

A ...

plasmid borne adjacent DNA segment and its homologous chromosomal counterpart resulting in the retention of the mutation in the chromosome, replacing the normal allele with the mutant one. The second method to introduce an altered allele into the chromosome employs gene conversion, described in Examples 37 and 43. In this method, an Ery Sac. erythraea strain carrying a deletion of a specified region of the eryA segment of the chromosome is used as a host. Into a Sac. erythraea multicopy plasmid that carries a selectable marker is cloned the wild type counterpart (segment 1) of the eryA segment mutant in the host. Subsequently, the desired homologous or heterologous DNA 10 segment to be introduced (segment 2) is cloned within the portion of segment 1 which is deleted in the mutant strain. The resulting plasmid is then introduced into the host employing selection for the marker. Among the transformants will be a population that have integrated segments 1 and 2 from the plasmid by the process of gene conversion 15 which can be verified by examination of the DNA among the colonies that have recovered the ability to produce erythromycin.

Two examples each of Types I, II and III alterations to the eryA DNA sequence and the resultant novel polyketides produced are described in the examples described herein. Examples 1 through 8, 9 through 12 and 13 20 through 16 describe the construction and effect of three Type I mutants. Examples 17 through 22 and 23 through 27 describe the construction of two Type II mutants and the effects of feeding two different synthetic substrates to the mutant strains. Examples 28 through 38 and 39 through 44 outline the steps in constructing Type III changes and their respective 25 effects on the structure of the novel polyketides produced. In Examples 1 through 7 a plasmid that contains a substantial deletion of the segment of the gene corresponding to the b-ketoreductase of module 5 is created, the altered gene is inserted into the Sac. erythraea chromosome to replace the wild type allele and the new strain carrying the altered gene is identified 30 and isolated. In Example 8, the new strain is fermented and the novel polyketide 5-oxo-5,6-dideoxy-3α-mycarosyl erythronolide B that results from the introduction of the mutant allele is isolated. In Examples 9 through 11, a mutation is introduced into the β -ketoreductase of module 2 and the mutated allele is then used to replace the wild type allele in the 35 chromosome. In Example 12, the strain carrying the altered allele is fermented and the novel compound 11-oxo-11-deoxyerythromycin A is isolated. Similarly, in Examples 13 through 16 a mutation is introduced





into the dehydratase of module 4 and the mutated allele is then used to replace the wild type allele in the chromosome. The strain carrying this altered allele is then fermented and the novel products 7hydroxyerythromycin A and 6-deoxy-7-hydroxyerythromycin A are 5 isolated. In Examples 17 through 21, a mutation is made in the DNA corresponding to the β-ketoacyl-ACP synthase of module 1 and introduced into the chromosome to replace the wild type allele. This mutation has the effect of arresting the synthesis of the polyketide chain and results in the Ery phenotype. The synthetic substrate (2S,3R,4S,5S)3,5-dihydroxy-2,4-10 dimethylhexanoic acid-ethyl ester is then made and fed to the mutant resulting in the production of the novel compound (14S,15S)14(1hydroxyethyl)erythromycin. Similarly, in Examples 22 through 24, a mutation is created in the β-ketoacyl-ACP synthase of module 2 and introduced into the chromosome to replace the wild type allele. In 15 Example 25 and 26, the synthetic substrate (2S,3S,4S,5S)2,4-dimethyl-3fluoro-5-hydroxyhexanoic acid-ethyl thioester is made and fed to the module 2 β-ketoacyl-ACP synthase mutant and the resulting novel compound 11-epifluoro-15-norerythromycin is isolated. In Examples 27 through 38, a copy of the DNA sequence corresponding to module 4 is 20 introduced into the deleted segment of the β-ketoacyl-ACP synthase of module 1 resulting in the production of the novel compound 14(1propyl)erythromycin. In Examples 40 through 44, a copy of the DNA sequence corresponding to module 5 is introduced into the deleted segment of the β -ketoacyl ACP synthase of module 1 resulting in the 25

GENERAL METHODS

production of the novel compound 14[1(1-hydroxypropyl)]erythromycin.

30 Materials, Plasmids and Bacterial Strains

35

Restriction endonucleases, T4 DNA ligase, nick-translation kit, competent E. coli DH5\alpha cells , X-gal, IPTG, and plasmids pUC19 and pUC12 are purchased from Bethesda Research Laboratories (BRL), Gaithersburg, MD. $[\alpha^{-32}P]dCTP$ and Hybond N are from Amersham Corp., Chicago, IL. Seakem LE agarose and Seaplaque low gelling temperature agarose are from FMC Bioproducts, Rockland, ME. E. coli K12 strains carrying the E. coli-Sac. shuttle plasmids pWHM3 or pWHM4 (Vara et al., J. Bacteriol., 171: 5872 (1989)) or the cosmids pS1 (Tuan et al., Gene, 90: 21 (1990)) and

Sac. erythraea strain NRRL2338 have been deposited in the culture collection of the Agricultural Research Laboratories, Peoria, IL and are available under the accession numbers NRRL XXXX, respectively. Staphylococcus aureus Th^R (thiostrepton resistant) is obtained by plating 10⁸ cells of S. aureus on agar medium containing 10 mg/ml thiostrepton and picking a survivor after 48 hr growth at 37°C. Thiostrepton is obtained from Squibb-Bristol Myers, New Brunswick, NJ. All other chemical and reagents are from standard commercial sources unless specified otherwise.

10

35

DNA Manipulations

Standard conditions (Maniatis et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982) are employed for restriction endonuclease digestion, agarose gel-electrophoresis, nick translation of DNA to make ³²P-labeled probes, 15 DNA ligation, and transformation of E. coli employing selection for ampicillin resistance (ApR) on LB agar plates. Plasmid DNA is isolated from minipreps of E. coli transformants by the boiling method (Maniatis et al., 1982, supra). DNA fragments are recovered from low melting agarose gels using the method of Langridge et al., 1980. Total DNA from 20 Sac. erythraea strains is prepared according to described procedures (Hopwood et al., Genetic Manipulation of Streptomyces, A Laboratory Manual, John Innes Foundation, Norwich, U.K., 1985). DNA is transferred from agarose gels onto Hybond N following the manufacturer's instructions. Hybridizations are performed in sealed bags 25 containing 10-20 ml of [1xNET (20xNET = 3 M NaCl, 0.3 M TrisHCl, 20 mM Na₂EDTA, pH 8.0), 5XDenhardt's solution (Maniatis et al., 1982, supra), 0.2 mg/ml denatured calf thymus DNA, 0.2% SDS, and 0.5-2x107 cpm of the nick-translated probe] for 16-20 hr at 65 °C. Filters are washed three times in 1xNET/0.1% SDS for 20 min each at room temperature, and once in 30 0.05xNET/0.1% SDS for 20 min at 70 °C. Filters are reused as described (Donadio et al., 1990).

Amplification of DNA fragments

Synthetic deoxyoligonucleotides are synthesized on an ABI Model 380A synthesizer (Applied Biosystems, Foster City, CA) following the manufacturer's recommendations. Amplification of DNA fragments is performed by the polymerase chain reaction (PCR) employing a Coy

PCT/US92/00427 WO 93/13663

ch primer, 1 µg of template thermocycler. Reactions contain 100 pmol c DNA (cosmid pS1 carrying the eryA segment from Sac. erythraea strain NRRL 2338), and 2.5 units of Thermus aquaticus DNA polymerase in a 100 ml volume of PCR buffer [50 mM KCl, 10 mM TrisHCl (pH 8.0) 2 mM MgCl₂, 0.01% gelatin) containing 200 mM of the 4 dNTPs. The above reagents are from Perkin Elmer Cetus, Norwalk, CT. The reaction mixture is overlaid with a drop of paraffin oil and subjected to 30-50 cycles. Each cycle consists of one 94 °C, one 55 °C and one 72 °C period, each of the duration of 3 min. The progress of the amplification is monitored by agarose gel-electrophoresis. The PCR primers described in the examples below are derived from the nucleotide sequence of eryA of FIG. 2.

Gene replacement and gene conversion

5

10

15

25

30

3.5

Protoplasts of Sac. erythraea strains are prepared and transformed with miniprep DNA isolated from E. coli according to published procedures (Yamamoto et al., 1986). Integrative transformants, in the case of pWHM3 derivatives, are selected after one round of non-selective growth of the primary ThR transformants as described by Weber et. al, Gene, 68: 173 (1988). Loss of the ThR phenotype is monitored by plating serial dilutions of a ThR integrant on non-selective medium, followed by 20 replica-plating on thiostrepton-containing medium. ThS (thiostreptonsensitive) colonies arise at a frequency of 10^{-2} (Donadio et al., 1990). The retention of the mutant allele is established by Southern hybridization of a few ThS colonies.

- A few hundred ThR colonies obtained by transformation of an eryA strain with pWHM4 derivatives are screened for antibiotic production by the agar-plug assay employing Staphylococcus aureus as ThR organism as described (Tuan et al., Gene, 90: 21 (1990)). The frequency of gene conversion between a 5 kb segment of homologous sequence and a strain carrying a small deletion is >25% (Tuan et al., Gene, 90: 21 (1990)). Colonies found to produce antibiotic activity are inoculated in SGGP (Yamamoto et al., 1986), protoplasts are prepared, and the regenerated protoplasts are scored for loss of the plasmid by replica-plating on non-selective medium. Th^S colonies are then rechecked for antibiotic production, and six producers are analyzed on Southern blots.

Fermentation

Sac. erythraea cells are inoculated into 100 ml SCM medium (1.5% soluble starch, 2.0% Soytone [Difco], 0.15% Yeast Extract [Difco], 0.01% CaCl₂) and allowed to grow at 32°C for 3 to 6 days. The entire culture is then inoculated into 10 liters of fresh SCM medium. The fermenter is operated for a period of 7 days at 32°C maintaining constant aeration and pH at 7.0. After fermentation is complete, the cells are removed by centrifugation at 4°C and the fermentation beer is kept in the cold until further use.

The present invention will now be illustrated, but is not intended to be limited, by the following examples:

Example 1 Construction of plasmid pABX9

15

10

The 9.6 kb <u>BamHI-XhoI</u> segment comprised between sequence coordinates 21.96 and 31.52 was isolated from cosmid pS1 and ligated to <u>SalI-digested pUC19 DNA</u>. The resulting mixture contained the desired plasmid pABX9.

20

Example 2 Construction of E. coli K12 DH5α/pABX9

Approximately 10 ng of plasmid pABX9, prepared as described in 2.5 Example 1, were transformed into <u>E</u>. <u>coli</u> K12 DH5α and a few of the resulting white Ap^R colonies that appeared on the LB-agar plates containing X-gal and ampicillin were analyzed for their plasmid content. One colony was found to carry pABX9, as verified by the observation of fragments of 3.93, 3.39, 2.01, 1.56, 0.87, and 0.48 kb in size upon agarose gel electrophoresis after <u>SmaI</u> digestion of the plasmid.

Example 3 Construction of plasmid pABX9DN

3 5 Plasmid pABX9, isolated from <u>E</u>. <u>coli</u> K12 DH5α/pABX9, was digested with <u>Ncol</u> and then treated with T4 DNA ligase. The resulting mixture contained the desired plasmid pABX9DN.

£

'n

Example 4 Construction of E. coli K12 DH5a/pABX9DN

Approximately 10 ng of plasmid pABX9DN, prepared as described in Example 3, were transformed into <u>E</u>. <u>coli</u> K12 DH5α and a few of the resulting white Ap^R colonies that appeared on the LB-agar plates containing X-gal and ampicillin were analyzed for their plasmid content. Colonies carrying pABX9DN exhibited a single <u>Nco</u>I fragment of 11.5 kb visible by agarose gel electrophoresis, confirming that the 813 bp <u>Nco</u>I - <u>Nco</u>I fragment from pABX9 has been deleted in pABX9DN.

Example 5 Construction of plasmid pABX95DN

Plasmid pABX95DN was digested with <u>Eco</u>RI and <u>Hin</u>dIII and ligated to pWHM3 digested with the same two enzymes. The resulting mixture contained the desired plasmid pABX95DN.

Example 6 Construction of E. coli K12 DH5α/pABX95DN

Approximately 10 ng of plasmid pABX95DN, prepared as described in Example 5, were transformed into E. coli K12 DH5a and a few of the resulting white Ap^R colonies that appeared on the LB-agar plates containing X-gal and ampicillin were analyzed for their plasmid content. Colonies carrying pABX95DN exhibited fragments of 8.8 and 7.2 kb visible in agarose gels after EcoRI and HindIII digestion.

Example 7 30 Construction of Sac. erythraea AKR5 carrying the eryAKR5 allele by gene replacement

Approximately 1 mg of plasmid pABX95DN, isolated from <u>E. coli</u> K12 DH5α/pABX95DN, was transformed into Sac. erythraea NRRL 2338 and stable Th^R colonies were isolated. Serial dilutions of one of these colonies were screened for the loss of the antibiotic resistance marker and total DNA from 5 Th^S colonies as well as from untransformed Sac. erythraea NRRL 2338 was digested with <u>Ss</u>tI and analyzed by Southern

hybridization employing the 0.8 kb <u>SalI</u> fragment between sequence coordinates 24.26 and 25.06 (from pABX9) as probe. Whereas NRRL 2338 showed one <u>SstI</u> band of 3.7 kb that hybridized to the probe, samples from four of the ThS strains exhibited a <u>SstI</u>-hybridizing band of 6.1 kb indicating the presence of the mutant allele. One of these colonies was kept and designated strain AKR5. It carries a deletion of 813 bp in the KR5 segment of *eryA* and is designated the <u>eryAKR5</u> allele.

Example 8

10 <u>Isolation, purification and properties of 5-oxo-5.6-dideoxy-3-a-mycarosyl</u> erythronolide B from Sac. erythraea AKR5

A 10-liter fermentation of Sac. erythrea AKR5 carrying the eryAKR5 allele in a Biolafitte fermentor using SNC Media. The fermentor was inoculated with 100 ml of a 3 day old seed. The pO2 was initially 80 ppm 15 and the temperature was maintained at 32°C. The pH was controlled to 7.0 ± 0.2 by addition of propionic acid or potassium hydroxide as needed. At harvest (3 days), the whole broth was extracted three times with 4-liter portions of ethylacetate. The combined extracts were concentrated under reduced pressure and the residue was chromatographed on a column (50 x 20 5 cm) of silica gel packed and loaded in toluene and eluted with a stepwise gradient of increasing concentration of isopropanol in toluene. Fractions were analyzed by TLC and spots were detected by spraying with anisaldehyde sulfuric acid spray reagent and heating. A major component giving blue colored spots eluted with approximately 7% isopropanol. 25 Fractions containing this material were combined and concentrated to a residue (800 mg). This was further chromatographed on a column (100 x 3 cm) of Sephadex LH-20 in chloroform-heptane-ethanol, 10:10:1, v/v/v. Fractions were analyzed as above, early fractions (9-13) yielded 5,6-dideoxy-3-a-mycarosyl-5-oxoerythronolide B (45 mg) which was crystallized from. 30 heptane/ethylacetate mixture to mp 163-164 °C.

CMR spectrum in CDCl3 (ppm downfield from TMS)

8.6	37.9	70.0
9.9	38.7	76.2
9.9	40.4	76.4
10.4	40.7	80.4
14.5	43.3	100.4
15.2	45.8	175.8
17.1	46.8	210.8
17.7	48.9	217.7
25.3	66.5	
25.5	69.4	

Structure was determined by single crystal X-ray diffraction.

Later fractions (15-17) yielded 5,6-dideoxy-5-oxoerythronolide B

(10 mg) and still later fractions yielded 5,6-dideoxy-6,6a-epoxy-5-oxoerythronolide B (2.8 mg).

1 0 Example 9 Construction of plasmid pALeryAKR2

contains the desired plasmid pALeryAKR2.

The 1.3 kb DNA segment comprised between coordinates 8.63-9.93

(fragment 1) is amplified by PCR employing two oligodeoxynucleotides, 1a

15 (5'-GGGAGCATGCTCTCGGTGCGCGGGCGGCGGCGC-3') and 1b (5'GCCCTGCAGCGCGTACTCCGAGGTGGCGGT-3'). Similarly, the 1.3 kb

DNA segment between coordinates 9.99-11.26 (fragment 2) is PCRamplified employing primers 2a (5'TGGTCTGCAGGCGAGGCCGGACACCGAGG-3') and 2b (5'
20 GGAAGAAGTCAAAGTTCCTCGGTCCCTTCT-3'). After digestion with

SphI + PstI (fragment 1) and PstI + EcoRI (fragment 2), the two fragments
are ligated to EcoRI + SphI-digested pWHM3. The resultant mixture

Example 10 Construction of E. coli K12 DH5a/pALeryAKR2

Approximately 10 ng of plasmid pALeryAKR2, prepared as described in Example 9, are transformed into Ε. coli K12 DH5α, and a few of the resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryA2KR2, 9.8 kb in size, and carrying a 2.6 kb EcoRI-SphI insert with an internal PstI site, is verified by SalI digestion (fragments at 2.91, 2.21, 1.61, 1.42, 1.08, 0.29, 0.12 and 0.10 kb are released, visible by agarose gel electrophoresis). pALeryAKR2 contains an in-frame deletion of 102 base pairs of the corresponding segment of the wild type eryA chromosomal DNA. The cloned segment in pALeryAKR2 is designated the eryAKR2 allele.

15

5

10

Example 11 Construction of Sac. erythraea AKR2 carrying the eryAKR2 allele by gene replacement

Approximately 1 mg of plasmid pALeryAKR2, isolated from E. coli 20 K12 DH5α/pALeryAKR2, is transformed into Sac. erythraea protoplasts and stable ThR colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and six ThS colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six ThS colonies and from untransformed Sac. 25 erythraea NRRL2338 is digested with PstI and with SalI and is then examined by Southern hybridization using the 2.6 kb EcoRI-SphI insert from pALeryAKR2 as probe. Whereas NRRL2338 contains a 39 kb PstI hybridizing band, colonies in which the mutation in KR2 has been introduced (strain AKR2) exhibit two bands of approximately equal 30 intensity, one at 27 kb and the other at 12 kb. The SalI digest, with bands at 1.04, 0.75, 0.29, 0.12 and 0.10 kb common to NRRL2338 and AKR2, but with the 1.16 kb band in NRRL2338 replaced by the 1.06 kb band in AKR2, confirms that the only change introduced into strain AKR2 is the deletion of the 102 bp segment from KR2, resulting in a strain carrying the 35 ervAKR2 allele.

,

Ŷ,

7

÷ .

Example 12 Isolation and purification of 11-deoxy-11-oxoerythromycin A

The fermentation beer of strain AKR2, cooled to 4°C is adjusted to pH 8.0 and is extracted sequentially with three equal volumes of methylene chloride. The combined methylene extracts are concentrated to an oily residue and partitioned between heptane and methanol. The methanol layer is removed, washed once with heptane and concentrated to a residue. The residue is digested in methylene chloride and washed once with potassium phosphate buffer pH 7.8 and once with water. The 10 methylene chloride layer is concentrated to a residue and digested in the lower phase (1:1:1, v/v/v) of a carbon tetrachloride; methanol; aqueous phosphate buffer (0.05 M, pH 7.0) system and chromatographed on an Ito Coil Planet Centrifuge in the same system. Fractions containing the desired 11-oxo-11-deoxyerythromycin A were combined, concentrated, 15 digested in methylene chloride, washed well with water and concentrated on rotary evaporator under reduced pressure to yield 11-deoxy-11oxoerythromycin A as an off-white solid froth. Its identity is confirmed by comparison with antibiotic L53-18A. 11-Deoxy-11-oxoerythromycin A is dissolved in tetrahydrofuran and the solution is diluted with an equal 20 volume of water. This is then acidified to pH 4.0 and allowed to stand at room temperature for 4 hours. The pH is adjusted to 9.0 and the solution is diluted with an equal volume of water and extracted with two volumes of methylene chloride. The combined methylene chloride extracts were evaporated to dryness under reduced pressure to yield antibiotic L53-18A 25 as a white solid.

Example 13 Construction of plasmid pALeryADH4

30

Primers 3a (GCGCGAGCTCGACGACCAGGGCGGCATGGT) and 3b (GGTGGCATGCTGCGACCACTGCGCGTCGGC) are used to PCR-amplify the 1.05 kb *eryA* segment of the <u>Sac. erythraea</u> chromosome between sequence coordinates 18.47-20.07 (fragment 3), and primers 4a

3 5 (AGCTGCATGCTCTGGACTGGGGACGGCTAG) and 4b (CGCGGGATCCCAGCTCCCACGCCGATACCG) are used to amplify the 1.35 kb segment between sequence coordinates 20.58-21.96 (fragment 4) as described in Example 1. Fragment 3 and 4, after digestion with <u>SstI + SphI</u>

and with <u>SphI</u> + <u>BamHI</u>, respectively, are ligated to <u>SstI</u> -, <u>BamHI</u>-digested pWHM3. The resulting ligation mixture contains the desired plasmid pALeryADH4.

Example 14 Construction of E. coli K12 DH5a/pALeryADH4

5

20

Approximately 10 ng of pALeryADH4, prepared as described in Example 13, are transformed transformed into Ε. coli K12 DH5α, and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryADH4, 9.6 kb in size, is verified by SphI + EcoRI digestion (fragments at 7.2, 1.35 and 1.05 kb are released). pALeryADH4 carries a 498 base pair in-frame deletion of the corresponding segment of the wild type eryA DNA. The cloned segment in pALeryADH4 is designated the eryADH4 allele.

Example 15 Construction of Sac. erythraea ADH4 carrying the eryADH4 allele by gene replacement

Approximately 1 mg of plasmid pALeryADH4, isolated from E. coli K12 DH5α/pALeryADH4, is used for transformation into Sac. erythraea protoplasts and stable ThR colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and 25 six ThS colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six ThS colonies and from untransformed Sac. erythraea NRRL2338 is digested with SphI and with SstI and examined by Southern hybridization using the 2.4 kb SstI-BamHI 30 insert from pALeryADH4 as probe. Strains in which the wild type allele has been replaced by the mutated copy show two SphI bands, one at 13.5 kb and the other at 12.4 kb, whereas the wild type strain exhibits a single band at 26 kb. The SstI pattern, with the 2.9 kb band from NRRL2338 being replaced in ADH4 by a 2.5 kb band, confirms that the 487 bp deletion created in plasmid pALeryADH4 has been transferred into the 35 chromosome of ADH4. Strains that carry the eryADH4 allele in place of the wild type sequence are designated Sac. erythraea ADH4.

Example 16

Isolation and characterization of 7-hydroxyerythromycin A and 6-deoxy-7hydroxyerythromycin A

The fermentation beer of strain ADH4 is cooled to 4°C and the pH is adjusted to 5.0. The mixture is extracted once with an equal volume of methylene chloride. The pH of the aqueous layer is readjusted to 9.0 and two further methylene chloride extracts are carried out. These two extracts are combined, washed with water and concentrated to a residue. This is 10 digested in 10 ml of the upper phase of a (3:7:5, v/v/v) mixture of hexane, ethylacetate, aqueous phosphate buffer (0.05 M, pH 7.5) and chromatographed on an Ito Coil Planet Centrifuge in the same system. Fractions containing the desired 7-hydroxyerythromycin were combined, concentrated, and partitioned between methylene chloride and dilute (pH 15 9.5) ammonium hydroxide solution. Fractions containing the desired 6deoxy-7-hydroxyerythromycin were combined, concentrated, and partitioned between methylene chloride and dilute (pH 9.5) ammonium hydroxide solution. The methylene chloride layers are washed with water and then concentrated to yield the desired 7-hydroxyerythromycin A and 20 6-deoxy-7-hydroxyerythromycin A as white foams.

Example 17 Construction of plasmid pALervAKS1

25 The 1.4 kb segment of eryA, between sequence coordinates 1.11-2.54 (fragment 5) and the 1.5 kb segment between sequence coordinates 2.88-4.37 (fragment 6) are PCR-amplified using primers 5a (TGCAGAATTCGCTGGCCGCGCTCTGGCGCT) and 5b (GAGAGCTGCAGCATGAGCCGCTGCTGCGGG), and 6a 30 (CATGCTGCAGGACTTCAGCCGGATGAACTC) and 6b (GAGGAAGCTTCCAGCCGGTCCAGTTCGTCC), respectively, as described in Example 9. After digestion with EcoRI + PstI (fragment 5) and PstI + <u>Hind</u>III (fragment 6), the two fragments are ligated to <u>Eco</u>RI + <u>Hind</u>IIIdigested pWHM3. The resulting mixture contains the desired plasmid 35 pALeryAKS1.

9

20

Example 18 Construction of E. coli K12 DH5a/pALeryAKS1

Approximately 10 ng of pALeryAKS1, prepared as described in

Example 17, are transformed into E. coli K12 DH5\(\alpha\), and a few of the resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAKS1, 10.1 kb in size, is verified by digestion with PstI + HindIII (fragments of 8.6 and 1.5 kb are observed by agarose gel

electrophoresis) and with SalI (fragments of 2.93, 2.21, 1.42, 1.37, 0.86, 0.54, 0.27, 0.14, 0.13, and 0.10 kb are observed). pALeryAKS1 carries an in-frame deletion of 282 base pairs of the corresponding wild type eryA DNA. The cloned insert in plasmid pALeryAKS1 is designated the eryAKS1 allele.

Example 19 Construction of Sac. erythraea AKS1 carrying the eryAKS1 allele by gene replacement

Approximately 1 mg of plasmid pALeryAKS1, isolated from E. coli K12 DH5α/pALeryAKS1, is used for transformation into Sac. erythraea 20 protoplasts and stable ThR colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and six ThS colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six ThS colonies and from untransformed Sac. erythraea NRRL2338 is digested with PstI and with 25 SmaI and examined in Southern hybridization employing the 2.9 kb EcoRI-HindII insert from pALeryAKS1 as probe. Colonies in which the wild type allele has been replaced by the mutated copy (strain AKS1) show two PstI bands, one at 34.5 and the other at 4.4 kb, whereas the wild type strain exhibits a single band at 39 kb. The SmaI pattern, with the 2.9 kb 30 band from NRRL2338 being replaced in AKS1 by a 2.6 kb band, confirms that the 282 bp created in plasmid pALeryAKS1 has been transferred into strain AKSI. Strains that carry the eryAKS1 allele are designated Sac. erythraea AKS1.

15

10

15

20

30

35

s

3

Example 20

Synthesis of (2S.3R.4S.5S)3.5-dihydroxy-2.4-dimethylhexanoic acid n-butyl thioester

A convenient source of this compound in chiral purity is the antibiotic oleandomycin. Oleandomycin (5 g) is dissolved in an aprotic solvent such as toluene and treated with diazabicyclo[5.4.0]undecene-5 (1 g) and heated for one hour. The resulting solution is poured into iced water, agitated well and the organic layer is drawn off and concentrated to a residue. The residue is digested in methylene chloride and treated exhaustively with a solution of ozone. The resulting ozonide is oxidatively decomposed with dilute hydrogen peroxide in sufficient aqueous ethanol to yield a monophasic mixture. This is further diluted with water and made 0.1 N with sodium hydroxide. The mixture is warmed for one hour at 70°C and then cooled before being acidified to pH 2.5 with dilute sulfuric acid. The mixture is then exhaustively extracted with methylene chloride. The combined extracts are concentrated to an oily residue and the desired lactone is recovered by chromatography on silica gel eluted with a gradient of tolueneisopropanol.

The δ -lactone is converted to the butyl thioester before feeding to Sac. erythrea AKS1 by refluxing with n-butylthiol in the presence of a catalytic amount of triethylamine.

25 <u>Example 21</u> Isolation of (14S,15S)14(1-hydroxyethyl)erythromycin A

The fermentation broth of AKS1 is cooled to 4°C and adjusted to pH 4.0 and extracted once with methylene chloride. The aqueous layer is readjusted to pH 9.0 and extracted twice with methylene chloride and the combined extracts are concentrated to a solid residue. This is digested in methanol and chromatographed over a column of Sephadex LH-20 in methanol. Fractions are tested for bioactivity against a sensitive organism, such as Staphylococcus aureus ThR, and active fractions are combined. The combined fractions are concentrated and the residue is digested in 10 ml of the upper phase of a solvent system consisting of n-heptane, benzene, acetone, isopropanol, 0.05 M, pH 7.0 aqueous phosphate buffer (5:10:3:2:5, v/v/v/v/v), and chromatographed on an Ito Coil Planet

بخ

€.

3

Centrifuge in the same system. Active fractions are combined, concentrated and partitioned between methylene chloride and dilute ammonium hydroxide (pH 9.0). The methylene chloride layer is separated and concentrated to yield the desired product as a white foam.

5

20

Example 22 Construction of plasmid pALeryAKS2

Primers 7a (CGCCCGAATTCGAGGCGCTGGGCGCCCGGAC) and 7b

10 (CCACCTGCAGCGCGGGACCTTCCAGCCCC), and primers 8a
(GTGGGTCGCTGCAGACGGTGACTGCGG) and 8b
(GGTCAAGCTTCGTCGGCGAGCAGCTTCTC) are used to PCR-amplify
the 1.45 kb eryA segment between sequence coordinates 5.71-7.16
(fragment 7) and the 1.5 kb eryA segment between sequence coordinates

15 7.22-8.70 (fragment 8), respectively. After digestion with EcoRI + PstI
(fragment 7) and with PstI + HindIII (fragment 8), the two fragments are
ligated to pWHM3 cut with EcoRI + HindIII. The resulting mixture
contains the desired plasmid pALeryAKS2.

Example 23 Construction of E. coli K12 DH5a/pALeryAKS2

Approximately 10 ng of pALeryAKS2, prepared as described in Example 22, are transformed into E. coli K12 DH5α, and a few of the resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAKS2, 10.1 kb in size, is verified by digestion with PstI + HindIII (fragments of 8.6 and 1.5 kb are observed by agarose gel electrophoresis) and with SstII (fragments of 4.0, 2.3, 2.0, 0.72, 0.43, 0.40, 0.20, 0.18, 0.13 and 0.11 kb observed). Plasmid pALeryAKS2 carries an inframe deletion of 60 base pairs of the corresponding wild type eryA DNA. This deletion removes the active site cysteine from KS2. The cloned insert in plasmid pALeryAKS2 is designated the eryAKS2 allele.

ŝ

설

Example 24

Construction of Sac. erythraea AKS2 carrying the eryAKS2 allele by gene replacement

Approximately 1 mg of plasmid pALeryAKS2, isolated from E. coli 5 K12 DH5α/pALeryAKS2, is used for transformation into Sac. erythraea protoplasts and stable ThR colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and six ThS colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six ThS colonies and from 10 untransformed Sac. erythraea NRRL2338 is digested with PstI and with SstII and examined in Southern hybridization employing the 2.9 kb EcoRI-HindII insert from pALeryAKS2 as probe. Colonies in which the wild type allele has been replaced by the mutated copy (strain AKS2) show two PstI 15 bands, one at 34.5 and the other at 4.4 kb, whereas the wild type strain exhibits a single band at 39 kb. The SstII pattern, with the 0.78 kb band from NRRL2338 being replaced in AKS2 by a 0.72 kb band, confirms that the 60 bp created in plasmid pALeryAKS2 has been transferred into strain AKS2. Strains that carry the ervAKS2 allele are designated Sac. erythraea 20 AKS2.

Example 25

Synthesis of (2R,3R,4S,5R)2,4-dimethyl-3-fluoro-5-hydroxyhexanoic acid n-butyl thioester

25

30

35

(2R,3S,4S,5R)3,5-Dihydroxy-2,4-dimethylhexanoic acid-δ-lactone (1 g) from Example 20 is digested in 10 ml of pyridine and treated with p-toluenesulfonyl chloride (1.3 g) and allowed to stand at room temperature overnight. The mixture is poured into iced water and extracted with methylene chloride and the methylene chloride is concentrated to the crude sulfonate ester. This is digested in acetonitrile (100 ml) and heated under reflux after the addition of tetrabutylammonium fluoride (1.75 g). After 6 hours the mixture is cooled, poured over iced water (300 ml) and extracted three times with 200 ml portions of methylene chloride. The combined methylene chloride extracts were concentrated and the residue was chromatographed on a column of silica gel eluted with a stepwise gradient of isopropanol (0 to 50%) in toluene. Fractions containing (2R,3R,4S,5R)2,4-dimethyl-3-fluoro-5-hydroxyhexanoic acid_d-lactone were

æ.

3

combined and concentrated to a white solid. The lactone is then converted to the n-butyl thiolester by refluxing in n-butyl thiol with a catalytic amount of triethylamine. Solvent is removed and the residue is digested in DMSO before feeding to fermentations of <u>Sac. erythraea</u> AKS2.

5

30

Example 26 Isolation and purification of 11-epifluoro-15-norerythromycin A

The fermentation broth of strain AKS2 is cooled to 4°C and adjusted to pH 4.0 and extracted once with ethylacetate. The aqueous layer is 10 adjusted to pH 9.0 and extracted twice with methylene chloride and the combined extracts are concentrated to a white solid. This is chromatographed over a column of Sephadex LH-20 in a mixture of heptane, chloroform, ethanol (10:10:1, v/v/v) and fractions containing the desired product are combined and concentrated to a solid residue. This is 15 further purified by countercurrent chromatography on an Ito Coil Planet Centrifuge on a system composed of carbon tetrachloride; methanol; 0.05 M; pH 7.0 aqueous potassium phosphate buffer (1:1:1, v/v/v). Fractions containing the desired 11-epifluoro-15-norerythromycin were combined, 20 and concentrated to a residue. This was digested in methylene chloride and dilute (pH 9.5) ammonium hydroxide and the methylene chloride layer was separated, washed with water and concentrated to yield the desired 11-epifluoro-15-norerythromycin A as white solid.

25 <u>Example 27</u> Construction of plasmid pALeryAM4.1

Primers 9a (GCGCCGAATTCTCGAGACGCGTGGGAGGCA) and 9b (TTGCGGTACCAGTAGGAGGCGTCCATCGCG) are employed to PCR-amplify the 2.0 kb *eryA* segment between sequence coordinates 17.35-19.38 (fragment 9). After digestion with <u>EcoRI + KpnI</u>, fragment 9 is ligated to pUC19 cut with the same two enzymes The resulting mixture contains the desired plasmid pALeryAM4.1.

Example 28 Construction of E. coli K12 DH5a/pALeryAM4.1

Approximately 10 ng of pALeryAM4.1, prepared as described in

Example 27, are transformed into <u>E. coli</u> K12 DH5a, and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.1, 4.7 kb in size, is verified by digestion with <u>SalI</u> (fragments of 2.8, 0.85, 0.53, 0.27 and 0.22 kb are observed by agarose gel electrophoresis).

Example 29 Construction of plasmid pALeryAM4.2

Primers 10a (GCTGGGATCCCGCGGGGGGGGGGGGGACAC) and 10b (CGGAACTCGGTGAGCATGCCGGGACTGCTC) are used to PCR-amplify the 2.1 kb *eryA* segment between sequence coordinates 21.94-24.00 (fragment 10). The 2.6 kb fragment KpnI(96)-BamHI(102) from cosmid clone pR1, and fragment 10 cut with BamHI + SphI, are ligated to pALeryAM4.1 cut with KpnI + SphI. The resulting mixture contains the desired plasmid pALeryAM4.2.

Example 30 Construction of E. coli K12 DH5a/pALeryAM4.2

25

30

*

Approximately 10 ng of pALeryAM4.2, prepared as described in Example 29, are transformed into <u>E. coli</u> K12 DH5a, and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.2, 9.3 kb in size, is verified by digestion with <u>XhoI</u> + <u>SphI</u> (to ensure that the entire 6.65 kb insert is released) and with <u>SalI</u>, with fragments of 2.8, 1.82, 1.09, 0.94, 0.85, 0.75, 0.45, 0.27, 0.22 and 0.13 kb are observed by agarose gel electrophoresis).

Example 31 Construction of plasmid pALeryAM1

The 2.9 kb <u>SmaI(4)-SmaI(20)</u> fragment from cosmid clone pRI is ligated to pUC12 cut with <u>SmaI</u>. The resulting mixture contains plasmid pALeryAM1.

10

20

25

30

35

Example 32 Construction of E. coli K12 DH5\alpha/pALeryAM1

æ

Ť

3

Approximately 10 ng of pALeryAM1, prepared as described in Example 31, are transformed into E. coli K12 DH5\alpha, and a few of the resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM1, 5.6 kb in size, is verified by digestion with SmaI (the 2.9 kb insert is realeased) and with SphI, with release of one 4.4 and one 1.07 kb bands. Both orientations of the insert in plasmid pALeryAM1 are useful.

Example 33 Construction of plasmid pALeryAM4.3

Plasmid pALeryAM1 is cut with <u>Xho</u>I to completion, partially with <u>SphI</u>, and the resulting 5.25 kb band, isolated from an agarose gel, is ligated to the 6.65 kb insert released from pALeryAM4.2 by <u>Xho</u>I + <u>SphI</u> digestion The resulting mixture contains the desired plasmid pALeryAM4.3.

Example 34 Construction of E. coli K12 DH5a/pALeryAM4.3

Approximately 10 ng of pALeryAM4.3, prepared as described in Example 33, are transformed into \underline{E} . \underline{coli} K12 DH5 α_z and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.1, 11.9 kb in size, is verified by $\underline{XhoI} + \underline{SphI}$ digestion (fragments of 6.65 and 5.25 kb are visible by agar se gelelectrophoresis). Plasmid pALeryAM4.3 carries the entire eryA module 4

10

30

3.5

*

inserted into the KS region f module 1. The cloned insert in pALeryAM4.3 is degnated the <u>eryAM412</u> allele.

Example 35 Construction of plasmid pALeryAM4.4

Plasmid pALeryAM4.3 is cut with <u>EcoRI + HindIII</u>, and the resulting 9.2 kb band, recovered from an agarose gel, is ligated to pWHM4 cut with the same two enzymes. The resulting mixture contains the desired plasmid pALeryAM4.4.

Example 36 Construction of E. coli K12 DH5\(\alpha\)/pALeryAM4.4

Approximately 10 ng of pALeryAM4.4, prepared as described in Example 35, are transformed into E. coli K12 DH5α, and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.1, 16.5 kb in size, is verified by EcoRI + HindIII digestion, with fragments of 9.2 and 7.3 kb released. Plasmid pALeryAM4.4 carries the eryAM412 allele on the Sac. erythraea multicopy vector pWHM4.

Example 37

25 <u>Construction of Sac. erythraea AM412 carrying the eryAM412 allele by gene conversion</u>

Approximately 1 mg of plasmid pALeryAM4.4, isolated from E. coli K12 DH5\(\alpha\)/pALeryAM4.4, is used for transformation into Sac. erythraea strain AKS1 protoplasts. A few hundred transformants are screened for antibiotic production by the agar-plug assay, and one of the colonies found to produce antimicrobial activity is cured of pALeryAM4.4 by protoplast formation and regeneration as described in General Methods. Total DNA from six antibiotic-producing, ThS colonies (strain AM412)and from strain AKS1 is digested with SphI and with XhoI and the resulting Southern blot is hybridized first to the 2.9 kb insert from pALeryAM1, and then to the 2.9 kb SstI(95)-SstI(101) fragment from plasmid pALeryAM4.2. With the first probe, the SphI band at 0.8 kb in strain AKS1 is seen to be replaced by a 7.5

æ

뾼

kb band in strain AM412, whereas the other two bands at 2.4 kb and 5.2 kb are unaffected. In the XhoI digest, the AKS1 band at 2.9 kb is replaced by a 9.6 kb band in AM412, with the other band at 5.2 kb conserved in both strains. Using the SstI(95)-SstI(101) fragment as probe, strain AKS1 exhibits one band at 25.5 kb and one at 17.9 kb in the SphI and XhoI digests, respectively, whereas, in addition to these bands, strain AM412 shows one SphI band at 7.5 kb and one XhoI band at 9.6 kb. In this way, it is established that the eryAKS1 allele has been converted into the eryAM412 allele in strain AM412.

10

Example 38 Isolation and purification of 14-(1-propyl)erythromycin A

At harvest the fermentation is adjusted to pH 9.5 and extracted twice with equal volumes of methylene chloride. The combined extracts 15 are washed once with water and concentrated to an oily residue. This is partitioned in a heptane methanol water (5:5:1, v/v/v) system and the lower layer is washed once with heptane and then concentrated to a semisolid residue. This is digested in methanol and chromatographed over a column of Sephadex LH-20 in methanol. Fractions are tested for 20 bioactivity in an agar diffusion assay on plates seeded with the macrolidesensitive strain Staphylococcus aureus ThR. Active fractions are combined and further purified by chromatography over silica gel a chloroform:methanol gradient containing 0.1% triethylamine. Fractions containing the desired 14-(1-propyl)erythromycin A are combined and 25 concentrated to yield the product as a white solid.

Example 39 Construction of plasmid pALeryAM5.1

30

The 4.7 kb eryA fragment between sequence coordinates 23.65-28.36 (fragment 11) is PCR-amplified employing primers 11a (ATGCTCGAGATCTCGTGGGAGGCGCTGGA) and 11b (AGAACTCGGTGAGCATGCCCGGGCCCGCCA). Fragment 11, after 3 5 digestion with XhoI + SphI, is ligated to the 5.25 kb fragment resulting from complete XhoI and partial SphI digestion of pALeryAM1, as in Example 33. The resulting mixture contains the desired plasmid pALeryAM5.1.

Example 40 Construction of E. coli K12 DH5\alpha/pALeryAM5.1

Approximately 10 ng of pALeryAM5.1, prepared as described in Example 39, are transformed into \underline{E} . coli K12 DH5 α_z and a few of the resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM5.1, 9.95 kb in size, is verified by SphI + XhoI 10 digestion, with fragments of 5.25 and 4.7 kb released, and by SmaI digestion where fragments of 3.39, 2.68 and 1.94 (doublet) kb are observed. Plasmid pALeryAM5.1 carries the entire eryA module 5 inserted into the β-ketoacyl ACP synthase region of module1. The cloned insert in plasmid pALeryAM5.1 is designated the eryA512 allele.

15

25

5

ś

Example 41 Construction of plasmid pALeryAM5.2

Plasmid pALeryAM5.1 is cut with EcoRI + HindIII and the resulting 20 6.3 kb fragment, recovered from an agarose gel, is ligated to pWHM4 cut with the same two enzymes. The resulting mixture contains the desired plasmid pALeryAM5.2.

Example 42 Construction of E. coli K12 DH5\alpha/pALeryAM5.2

Approximately 10 ng of pALeryAM5.2, prepared as described in Example 41, are transformed into E. coli K12 DH5\alpha, and a few of the resulting white ApR colonies that appear on the LB-agar plates containing 30 X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM5.2, 13.6 kb in size, is verified by digestion with EcoRI + HindIII, with fragments of 7.3 and 6.3 kb released. Plasmid pALeryAM5.2 contains the ervAM512 allele on the Sac. ervthraea multicopy vector pWHM4.

35

Example 43 Construction of Sac. erythraea AM512 carrying the eryAM512 allele by gene conversion

خ

Approximately 1 mg of plasmid pALeryAM5.2, isolated from E. coli 5 K12 DH5 α /pALeryAM5.2, is used for transformation into <u>Sac</u>. <u>erythraea</u> strain AKS1 protoplasts. A few hundred transformants are screened for antibiotic production by the agar-plug assay, and one of the colonies found to produce antimicrobial activity is cured of pALeryAM5.2 by protoplast formation and regeneration as described in General Methods. Total DNA 10 from six antibiotic-producing, ThS colonies (strain AM512)and from strain AKS1 is digested with SphI and with XhoI and the resulting Southern blot is hybridized first to the 2.9 kb insert from pALeryAM1, and then to the 0.8 kb NcoI(119)-NcoI(123) fragment from plasmid pALeryAM5.1. With the first probe, the SphI band at 0.8 kb in strain AKS1 is replaced by a 5.5 kb 15 band in strain AM512, whereas the other two bands at 2.4 kb and 5.2 kb are unaffected. In the XhoI digest, the AKS1 band at 2.9 kb is replaced by a 7.6 kb band in AM512, with the other band at 5.2 kb conserved in both strains. Using the NcoI(119)-NcoI(123) fragment as probe, strain AKS1 exhibits one band at 25.5 kb and one at 17.9 kb in the SphI and XhoI digests, 20 respectively, whereas, in addition to these bands, strain AM512 shows one SphI band at 5.5 kb and one XhoI band at 7.6 kb. In this way, it is established that the eryAKS1 allele has been converted into the eryAM512 allele in strain AM512.

25

30

Example 44 Isolation and purification of 14[1(1-hydroxypropyl)]erythromycin A

At harvest the pH of the fermentation of AM512 is adjusted to 9.5 and the mixture is extracted twice with equal volumes of ethylacetate. The combined ethylacetate extracts are washed with water, dried and partitioned in a heptane, methanol, water (5:5:1, v/v/v) system. The lower (methanolic phase) is washed with an equal volume of heptane and is concentrated to a residue. This is chromatographed on a Sephadex LH-20 column in methanol and fractions containing the desired 14[1(1-hydroxypropyl)]erythromycin A are concentrated and further purified by chromatography on an Ito Coil Planet Centrifuge in a system consisting of n-heptane, benzene, acetone, isopropanol, 0.65 M, pH 7.0 aqueous

potassium phosphate buffer (5:10:2:3:5, v/v/v/v). Fractions containing the desired product are concentrated to a solid residue and partitioned between methylene chloride and dilute (pH 9.5) ammonium hydroxide. The organic layer is washed with water and concentrated to yield 14[1(1-hydroxypropyl)]erythromycin A as a white solid.

Although the present invention is described in the Examples listed

5

above in terms of preferred embodiments, they are not to be regarded as limiting the scope of the invention. The above descriptions serve to 10 illustrate the principles and methodologies involved in creating the three types of mutations that can be introduced into the eryA segment of the Sac. erythraea chromosome that result in the synthesis of novel polyketide products. Although single Type I alterations, leading to the production of 5-oxo-5,6,-dideoxy-3α-mycorosyl erythronolide B, 11-oxo-11deoxyerythromycin A, 7-hydroxyerythromycin A, 7-oxo-15 7deoxyerythromycin A, 5-desosaminyl-3-oxo-3-deoxyerythronolide A, and Δ-6,7-anhydro-6-deoxyerythromycin A are specified herein, it is obvious that other Type I changes can be introduced into the eryA segment leading to novel polyketide structures. Among the additional Type I 20 alterations that can be obtained are those in which two or more modules are affected leading to the synthesis of novel polyketides. Examples of combinations of two Type I alterations leading to useful compounds include but are not limited to: mutants of the the β-ketoreductase of module 2 (KR2) and the β-ketoreductase of module 4 (KR4) leading to the 25 formation of 7,11-dioxo-7,11-dideoxyerythromycin A; mutants of KR2 and the β-ketoreductase of module 6 (KR6) leading to the formation of 3,11dioxo-3,11-dideoxy-5-desosaminylerythronolide A; mutants of KR2 and the dehydratase of module 4 (DH4) leading to the synthesis of 7-hydroxy-11-oxo-11-deoxyerythromycin A; mutants of KR2 and the enoylreductase 30 of module 4 (ER4) leading to the synthesis of Δ -6,7-anhydro-11-oxo-11deoxyerythromycin A; mutants of KR4 and KR6 leading to the synthesis of 3,7-dioxo-3,7-dideoxy-5-desosaminylerythronolide A; mutants of KR6 and DH4 leading to the synthesis of 3-oxo-3-deoxy-5-desosaminyl-7hydroxyerythronolide A; mutants of KR6 and ER4 leading to the synthesis 35 of 3-oxo-3-deoxy-5-desosaminyl- Δ -6,7-anhydroerythronolide A. Examples of combinations of three Type I alterations leading t the synthesis f novel polyketides include but are not limited to: mutants of KR2, KR4 and KR6 leading to the synthesis of 3,7,11-trioxo-3,7,11-trideoxy-5-

PCT/US92/00427 WO 93/13663

32

desosaminylerythronolide A; mutants of KR2, KR6 and DH4 leading to the synthesis of 3,11-dioxo-3,11-dideoxy-5-desosaminyl-7hydroxyerythronolide A; mutants of KR2, KR6 and ER4 leading to the synthesis of 3,11-dioxo-3,11-dideoxy-5-desosaminyl-D-6,7-

anhydroerythronolide A. All combinations of two or three Type I mutants, the Sac. erythraea strains that carry said combinations and the corresponding polyketides produced from said strains, therefore, are included within the scope of the present invention.

Although the Type II mutants specified herein have been constructed in the β -ketoacyl ACP synthase of module 1 (KS1) and the β -10 ketoacyl ACP synthase of module 2 (KS2), other Type II mutants can be constructed in other domains to result in the synthesis of novel polyketide structures upon feeding with appropriate substrate analogs. Other Type II mutants include but are not limited to: inactivation of the either of the acyltransferases or acyl carrier proteins of module 1, or the 15 acyltransferase or acyl carrier protein of module 2, the β -ketoacyl ACP synthase, acyltransferase or acyl carrier protein of module 3, module 4 or module 5. Furthermore, compounds other than (2S,3R,4S,5S)3,5dihydroxy-2,4-dimethylhexanoic acid-ethyl thioester and (2S,3S,4S,5S)2,4-20 dimethyl-3-fluoro-5-hydroxyhexanoic acid-ethyl thioester specified herein can be synthesized and fed to strains AKS1 or AKS2 specified herein or other strains that carry other Type II mutations to result in the creation of novel polyketides that are within the scope of the present invention.

Although two examples of Type III alterations are specified herein, it is apparent to those skilled in the art that many other examples of Type III changes are possible. Strains of Sac. erythraea carrying changes of this type offer the very high potential for the production of novel polyketides of specified structure, since they do not require synthetic substrates as do Type II mutants and they are not limited to the formation of derivatives 30 of erythromycin, as in the case of Type I mutants. In the embodiments of Type III mutants specified herein, we have illustrated how a second copy of a complete module can be introduced at a desired position by gene conversion to result in the synthesis of 14-(1-propyl)erythromycin A or 14-[1(1-hydroxypropyl])erythromycin A. These alterations make use of the high conservation and simultaneous lack of specificity of the β-ketoacyl ACP synthases of modules 1 and 2, thereby making possible the construction of hybrid β -ketoacyl ACP synthase functions consisting of portions of proteins derived from different modules. Those skilled in the

3

25

3.5

10

15

20

25

30

35

3

art understand, therefore, that it is possible, as exemplified for KS1 and KS2, to delete a small portion of the β -ketoacyl ACP synthase of other modules and to construct strains carrying such alterations which can then be employed as hosts for introducing at the deleted β -ketoacyl ACP synthase location a second copy of any homologous module. Furthermore, as exemplified herein, it is also possible to delete any segment of eryA by ligation of two non-contiguous PCR-generated fragments and to subsequently construct strains, therefore, devoid of any or all portions of any module. Such strains deleted of a full module can be employed for reintroduction of either the same or a different module at a different location. It is possible, therefore, to determine the novel structures desired and then create a series of Sac. erythraea strains containing the corresponding arrangements of eryA modules that would produce said novel structures that are included within the scope of the present invention. Additional examples of novel compounds produced from the construction of Type III alterations include but are not limited to 11-deoxyerythromycin, resulting from the insertion of the eryA segment encoding DH4 and ER4 in module 2.

Moreover, it will also be apparent that two or more modules can be excised and introduced into various sites of the Sac. erythraea chromosome to produce novel polyketides of predicted structure such as the introduction of the eryA segment encoding DH4 and ER4 in both module 1 and module 2 to result in the production of 14(R)[1-hydroxypropyl]11-deoxyerythromycin A. All combinations, therefore, of Type III alterations and the strains of Sac. erythraea that carry said alterations as well as the polyketides produced from said strains are included within the scope of the present invention.

In addition, it is also possible to create combinations of Type I, Type II and Type III alterations and insert such alterations into Sac. erythraea to produce novel polyketides. Examples of such combinations include but are not limited to the following. The combination of a Type I alteration, such as an alteration in DH4 and a Type II alteration, such as a mutation in the KS1 to result in the formation of (14S,15S)14-[1-hydroxyethyl]-7-hydroxyerythromycin A when the strain of Sac. erythraea carrying such alterations is fed with the compound (2S,3R,4S,5S)3,5-dihydroxy-2,4-dimethylhexanoic acid ethyl ester. The combination of a Type I alteration, such as an alteration in DH4 and a Type III alteration, such as found in Sac. erythraea strain AM412, wherein a copy of the DNA segment of

٤

10

15

20

25

module 4 is introduced in module 1, such that the Sac. erythraea strain so constructed produces the compound 7-hydroxy-14-propylerythromycin A. All combinations of two or more alterations of Type I, Type II and Type III alterations, the Sac. erythraea strains that carry such alterations, and the polyketides produced from such strains are included within the scope of the present invention. It will also occur to those skilled in the art that novel structures can be produced by altering the specificity of the acyltransferase functions in any module. Examples include: replacement of the acyltransferase domains of modules 1, 2, 3, 4, 5, or 6 in eryA with those of modules 4, 4, 2, 2, 2, and 4, respectively, to result in the production of 12-epierythromycin A,10-epierythromycin A, 8epierythromycin A, 6-epierythromycin A, 4-epierythromycin A and 2epierythromycin A, respectively, that are included within the scope of the present invention.

It should be emphasized that the introduction of an entire eryA module at a different location, as exemplified for the construction of Sac. erythraea strains AM412 and AM512 in Examples 29 and 35, respectively, does not rely on homologous recombination between the incoming eryA module and the host chromosome. Rather, gene conversion of the host allele with the eryA allele residing on the multicopy plasmid requires DNA sequences homologous to the host allele flanking the incoming module. Thus, any module carrying the desired specificities, either from homologous or heterologous sources, can be employed in gene conversion of the host allele, provided that is flanked by segments of homology. It will occur to those skilled in the art, therefore, that, given the large number of natural polyketide molecules existing, a wide variety of additional novel molecules of predicted structure can be produced in Type III mutants containing an additional module of desired specificities or where an endogenous module is replaced by an exogenous one. The length of the acyl chain can be easily controlled by suitably changing the 30 number of modules involved in its synthesis. Similarly, the introduction of keto, hydroxy, enoyl, or methylene groups at specific points along the acyl chain can be easily achieved by introducing the proper b-carbon processing functions (β-ketoreductase, dehydratase and enoylreductase) in the required modules. Exogenous modules constitute the source of 35 specificities for starter and extender units other than those employed by Sac. erythraea for erythromycin biosynthesis, making it thereby possible to employ, for example, malonylCoA or (2R)- or (2S)ethylmalonylCoA, etc.

3

as extender units, and acetyl CoA, butyryl CoA, etc. as the starter unit. The result will be the formation of erythromycin analogs containing the desired functional groups and side chains with the desired stereochemistry. As an extension of the examples reported with eryA, the construction of a Sac. erythraea strain carrying a heterologous module inserted into eryA requires: (i) cloning of the genes from any other Actinomyces producing a polyketide with desired structural features; (ii) mapping of the modular organization of the cloned genes by low stringency hybridization and restriction analysis; (iii) locating the module carrying the desired specificities by partial sequencing; (iv) precise excision 10 of the desired genetic element and cloning into a vector suitable for gene conversion; (v) construction and transformation of a Sac. erythraea strain suitable for gene conversion and screening for the novel compound. Any module, or portion thereof, can thus be precisely excised from the genome of a polyketide-producing microorganism and introduced into suitable Sac. erythraea strains to create a novel polyketide of predicted structure. Thus, replacement of the acyltransferase segments of modules 1, 2, 3, 4, 5,or 6 in eryA with the acyltransferase segment specific for malonyl CoA, such as can be found in the polyketide synthase genes for the synthesis of pikromycin in Streptomyces venezuelae, to result in the synthesis of 12-20 norerythromycin A, 10-norerythromycin A, 8-norerythromycin A, 6norerythromycin A, 4-norerythromycin A and 2-norerythromycin A, respectively, that are included within the scope of the present invention. In addition, replacement of the acyltransferase segments of modules 1, 2, 3, 25 4, 5, or 6 in eryA with an acyltransferase specific for (2R)-ethylmalonyl CoA, such as can be found in the polyketide synthase genes for the synthesis of spiramycin in Streptomyces ambofasciens, will result in the formation of 12-homoerythromycin A, 10-homoerythromycin A, 8epihomoerythromycin A, 6-epihomoerythromycin A, 4epihomoerythromycin A and 2-homoerythromycin A, respectively, all of which are included within the scope of the present invention. Similarly, introduction of acyltransferase segments carrying desired specificities for

30 the starter or extender unit into eryA DNA that results in the synthesis of novel compounds are included within the scope of the present invention.

35 The erythromycin analogs produced by the method of this invention are structurally similar to known antibacterial and prokinetic agents.

It will also occur to those skilled in the art that genetic manipulations described herein need not be limited to Sac. erythraea.

ş

Ţ.

20

Suitable hosts are any other polyketide-producing Actinomyces where DNA can be precisely inserted into the chromosome. Hence, the choice of a convenient host is based solely on the relatedness of the novel polyketide to a natural counterpart so as to minimize the number of module rearrangements required for its biosynthesis. Therefore, Type I, 5 Type II and Type III alterations can be constructed in other Actinomyces employing either endogenous or exogenous modules to produce novel polyketides employing strategies analogous to those described herein for Sac. erythraea. Thus all Type I, Type II or Type III mutations or various combinations thereof constructed in other actinomycetes according to the principles described herein, and the respective polyketides produced from such strains, are included within the scope of the present invention. Examples of polyketides that can be altered by creating Type I, Type II or Type III changes in the producing microorganisms include, but are not limited to macrolide antibiotics such as erythromycin, tylosin, spiramycin, 15 etc.; ansamacrolides such as rifamycins, maytansines, etc.; polyketide antibiotics such as tetracycline; polyethers such as monesin, salinomycin, etc.; polyenes such as candicidin, amphothericins; immunosuppressants such as FK506, ascomycin, rapamycin, etc. and other complex polyketides such as avermectin.

Whereas the novel derivatives or modifications of erythromycin described herein have been specified as the A derivatives, such as 7hydroxyerythromycin A, 11-oxo-11-deoxyerythromycin A, 14[1(1hydroxypropyl)]erythromycin A, etc., those skilled in the art understand 25 that the wild type strain of Sac. erythraea produces a family of erythromycin compounds, including erythromycin A, erythromycin B, erythromycin C and erythromycin D. Thus, modified strains of Sac. erythraea, such as strain AKR2, for example, would be expected to produce the corresponding members of the 11-oxo-11-deoxyerythromycin family, including 11-oxo-11-deoxyerythromycin A, 11-oxo-11-deoxyerythromycin 30 B, 11-oxo-11-deoxyerythromycin C, and 11-oxo-11-deoxyerythromycin D. Similarly, strain AM412 would be expected to produce not only 14(1propyl)erythromycin A but also the other members of the 14(1propyl)erythromycin family including 14(1-propyl)erythromycin B, 14(1-35 propyl)erythromycin C and 14(1-propyl)erythromycin D. Similarly, all other modified strains of Sac. erythraea described herein that produce novel erythromycin derivatives would be expected to produce the A, B, C, and D forms of said derivatives. Therefore, all members of the family of

WO 93/13663 PCT/US92/00427

37

each of the novel polyketides described herein are included within the scope of the present invention.

5

10

15

20

25

30

35

8

3,

45 ·

Variations and modifications of the methods for obtaining the desired plasmids, hosts for cloning and choices of vectors and segments of eryA DNA to clone and modify, other than those described herein that result in substantially the same strains and same products as those described herein will occur to those skilled in the art. For example, although we have described the use of the plasmids pWH3 and pWHM4 as E. coli-Sac. erythraea shuttle vectors, other vectors can be employed wherein all or part of pWHM3 or pWHM4 is replaced by other DNA segments that function in a similar manner, such as replacing the pUC19 component of pWHM3 and pWHM4 with pBR322, available from BRL, employing different segments of the pIJ101 or pJV1 replicons in pWHM3 and pWHM4, respectively, or employing selectable markers other than thiostrepton- and ampicillin-resistance. These are just few of a long list of possible examples all of which are included within the scope of the present invention. Similarly, the segments of the eryA locus subcloned into pWHM3 for generating strains AKS1, AKS2, etc. specified herein can readily be substituted for other segments of different length encoding the same functions, either produced by PCR-amplification of genomic DNA or of an isolated clone, or by isolating suitable restriction fragments from Sac. erythraea. In the same way, it is possible to create eryA strains carrying mutations functionally equivalent to those described herein by deleting different portions of the corresponding genes, by creating insertions into them, or by site-directed mutagenesis of specific nucleotide residues. Moreover, Sac. erythraea strains with mutant alleles other than the β ketoacyl ACP synthase portions of eryA can be employed as hosts for gene conversion; Type III mutants can be constructed by double reciprocal crossover as exemplified for Type I and Type II mutants rather than by the gene conversion method described herein. Additional modifications include changes in the restriction sites used for cloning or in the general methodologies described above. All such changes are included in the scope of the invention. It will also occur to those skilled in the art that different methods are available to ferment Sac. erythraea, to extract the novel polyketides specified herein, and to synthesize substrate analogs, and that all such methods are also included within the scope of the present invention.

WO 93/13663 PCT/US92/00427

38

It will be apparent that many modifications and variations of the invention as set forth herein are possible without departing from the spirit and scope thereof, and that, accordingly, such limitations are imposed only as indicated by the appended claims.

5

સ

É

Ç

Ł

PCT/US92/00427

ņ

25

30

What is claimed is:

- 1. A method for directing the biosynthesis of specific polyketide analogs by genetic manipulation of a polyketide-producing microorganism, said method comprising the steps of:
- (1) isolating a polyketide biosynthetic gene-containing DNA sequence;
- (2) identifying enzymatic activities associated within said genecontaining DNA sequence;
- 10 (3) introducing one or more specified changes into said genecontaining DNA sequence which codes for one of said enzymatic activities resulting in an altered DNA sequence;
 - (4) introducing said altered DNA sequence into a polyketideproducing microorganism to replace the original sequence;
- 15 (5) growing a culture of the altered microorganism under conditions suitable for the formation of the specific polyketide analog; and (6) isolating said specific polyketide analog from the culture.
- 2. The method of claim 1 wherein said polyketide biosynthetic gene-20 containing DNA sequence comprises genes which encode the enzymatic activities comprising a polyketide synthase.
 - 3. The method of claim 2, wherein said polyketide synthase enzymatic activities comprise β -ketoreductase, dehydratase, acyl carrier protein, enoylreductase, β -ketoacyl ACP synthase, and acyltransferase.
 - 4. The method of claim 1 wherein said alteration which occurs in the DNA sequence results in the inactivation of one or more enzymatic activities involved in the processing of the β -carbonyl of said polyketide.
 - 5. The method of claim 4, wherein said inactivated enzymatic activities affecting the processing of the β -carbonyl of said polyketide comprise β -ketoreductase, dehydratase, and enoylreductase.
- 3 5 6. The method of claim 4 wherein said alteration in the DNA sequence results in the addition of one or more enzymatic activities involved in the β-carbonyl processing of said polyketide.

PCT/US92/00427

3

30

35

- 7. The method of claim 6 wherein said additional enzymatic activities are selected from the group consisting of β -ketoreductase, β -ketoreductase and dehydratase, and β -ketoreductase, dehydratase and enoylreductase.
- 5 8. The method of claim 1 wherein said alteration occurring in the DNA segment results in the inactivation of one or more enzymatic activities involved in the condensation of carbon units to the nascent polyketide structure.
- 10 9. The method of claim 8 wherein said enzymatic activities affecting the condensation of carbon units to the nascent polyketide structure comprise β-ketoacyl ACP synthase, acyl carrier protein and acyltransferase.
- 10. The method of claim 1 wherein said alteration in the DNA
 15 sequence results in the change of the length of the polyketide synthesized.
 - 11. The method of claim 10 wherein said alteration results in the increase of the length of the polyketide.
- 20 12. The method of claim 11 wherein said alteration comprises the addition of DNA sequences encoding the enzymatic activities consisting of acyltransferase, acyl carrier protein and β-ketoacyl ACP synthase.
- 13. The method of claim 10 wherein said alteration results in the25 decrease of the length of the polyketide.
 - 14. The method of claim 13 wherein said alteration consists of the deletion of a DNA segment between two sequences encoding corresponding enzymatic activities.
 - 15. The method of claim 14 wherein said corresponding enzymatic activities are selected from the group consisting of β -ketoreductases, dehydratases, acyl carrier proteins, enoylreductases, β -ketoacyl ACP synthases, and acyltransferases.
 - 16. The method of claim 1 wherein said alteration consists in the replacement of the DNA segment encoding an acyltransferase with a DNA segment encoding an acyltransferase of different specificity.

۴,

ś

15

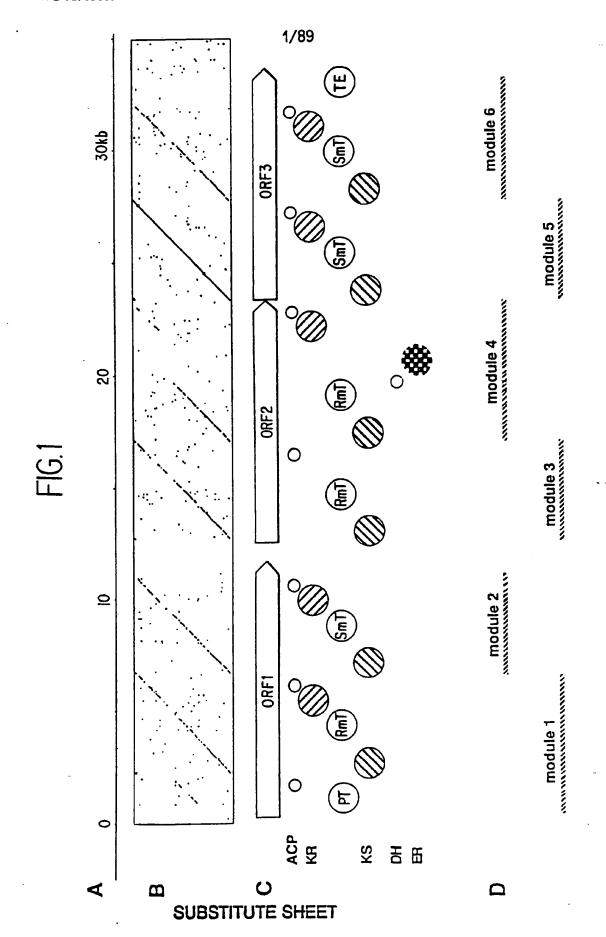
- 17. The method of claim 1 wherein said DNA sequence is isolated from a species from the Actinomycetales family.
- 5 18. The method of claim 17 wherein said DNA sequence is isolated from a genus selected from the group consisting of Actinomyces,

 Dactylosporangium, Micromonospora, Nocardia, Sac., Streptoverticillium, and Streptomyces.
- 10 19. The method of claim 17 wherein said genus is selected from the group consisting of Saccharapolyspora and Streptomyces.
 - 20. The method of claim 19 wherein said genus is Saccharapolyspora and the species is erythraea.
- 21. The method of claim 19 wherein said genus is Streptomyces and the species is hygroscopicus.
- 22. The method of claim 1 wherein said polyketide is selected from the group consisting of macrolides, tetracyclines, polyethers, polyenes, ansaymcins and derivatives or analogs thereof.
 - 23. The method of claim 22 wherein said polyketide is a macrolide.
- 25 24. The method of claim 23 wherein said macrolide is an erythromycin.
 - 25. The method of claim 24 wherein said erythromycin analog is selected from the group consisting of 11-oxo-11-deoxyerythromycin A, 7-hydroxyerythromycin A, 6-deoxy-7-hydroxyerythromycin A, 7-
- 30 oxoerythromycin A, 3-oxo-3-deoxy-5-desosaminylerythronolide A, Δ-6,7-. anhydroerythromycin A, ((14S,15S)14(1-hydroxyethyl)erythromycin A, 11-epifluoro-15-norerythromycinA, 14-(1-propyl)erythromycin A, 14(1-propyl)erythromycin A, and 14[1(1-hydroxypropyl)]erythromycin A.
- 3 5 26. The method of claim 1 wherein said DNA sequence, designated eryA, encodes the enzymatic activities associated with the formation of 6-deoxyerythronolide B.

PCT/US92/00427

- 27. The method of claim 26 wherein said DNA sequence comprises: the DNA sequence of Figure 2.
- 28. The method of claim 1 wherein said gene-containing DNA sequence encodes one or more enzymatic activities in the rapamycin biosynthetic pathway.
 - 29. The method of claim 23 wherein said macrolide is a rapamycin analog.

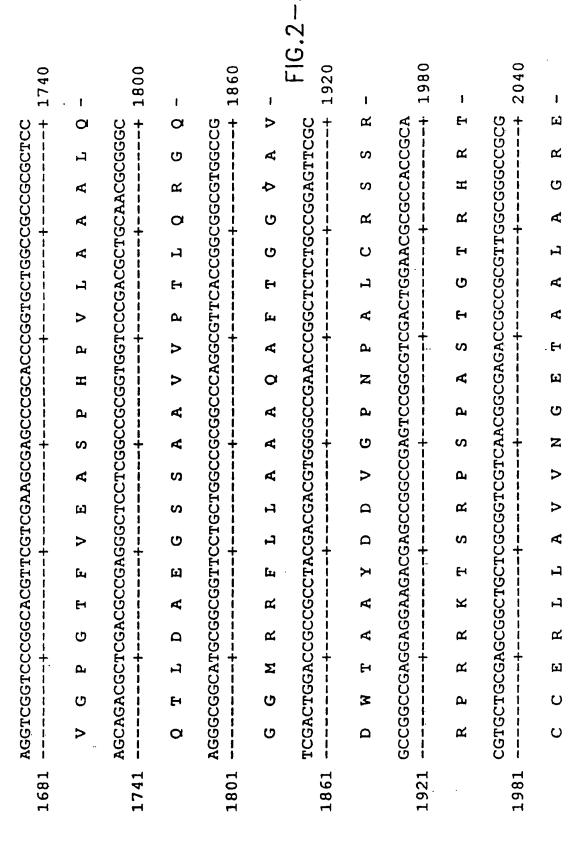
30. A compound selected from the group consisting of 7-hydroxyerythromycin A; 6-deoxy-7-hydroxyerythromycin A; 7-oxoerythromycin A, 3-oxo-3-deoxy-5-desosaminyl-erythronolide A; Δ-6,7-anhydroerythromycin A; ((14S,15S)14(1-hydroxyethyl)erythromycin A; 11-epifluoro-15-norerythromycin A; 14-(1-propyl)erythromycin A; 14(1-propyl)erythromycin A; and 14[1(1-hydroxypropyl)]erythromycin A.



a

						F1G.2					
	080		140		1200		1260	1	1320	1	1380
GESRVFAAMDACARAFE-	AGCCCGTGACCGACTGGACGCTGGCGCCAGGTCCTGGACTCTCCCGAGCAGTCGCGCCGCGCG	T D W T L A Q V L D S P E Q S R R V-	rcgaggicgiccagcccgcccrgrrcgcggrgcagacgrcgcrggccgcgcrrggcgcrg+ 1140	V Q P A L F A V Q T S L A A L W R S -	CCTTCGGCGTGACCCCCGACGCCGTGGTGGGCCACAGCATCGGCGAGCTGGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC	VTPDAVVGHSIGELAAAH-	ACGTGTGCGGTGCGGCCGGTGCCGCCGCCGCGCGCGCGCCGC	GAAGAADARAARUWSRE-	AGATGATTCCGTTGGTGGGCAACGGCGACATGGCAGCCGTCGCGCTCTCCGCCGCGGACGAGA +++++++	PLVGNGDMAAVALSADEI-	TCGAGCCGCGCATCGCCCGGTGGGACGACGACGTGGTGCTGGCCGGGGTCAACGGTCCGC
1	AGCCCC	Ь	TCGAGO	ы С	CCTTCC	Fi C)	ACGTG1	o	AGATG!	Σ	TCGAGC
	1021		1081		1141		1201		1261		1321

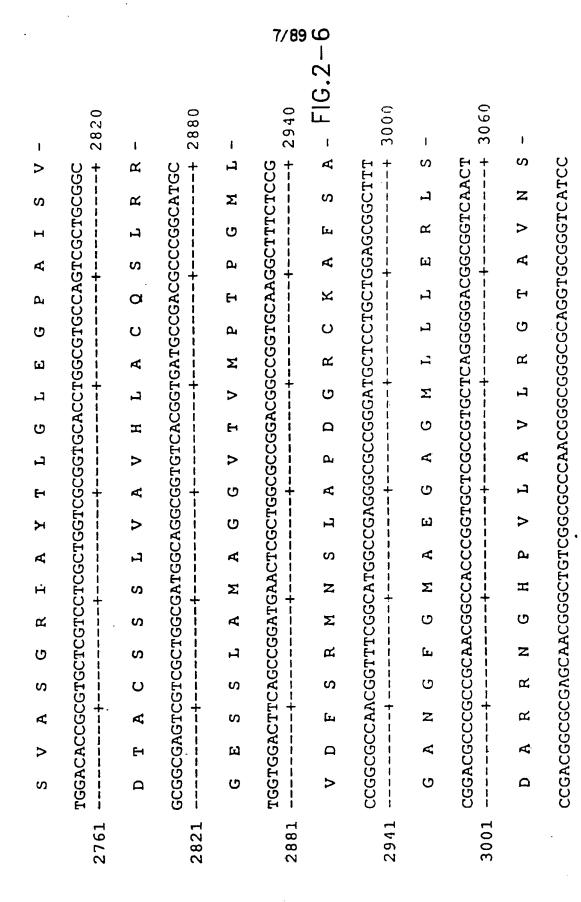
					3/89	-2				
						FIG.2-2				
	0		0		0.9	FIG	0		30	
ı	1440	ı	1500	i	1560	ı	1620	1	1680	i
ሺ	99	æ	50 + +	Q	99	>	CH †	×	99 +	ចា
വ	OHO I	လ	AGGTCG	>	CGGAGG	ഥ	CGA	Ω	CCI	H
ග	3CT	ы	GCA	ø	CIC	S	299	æ	000	Æ
z	3GA(េ	3GC +	æ	TGG +	G	GGT +	>	GTC	S
გ > ს	CCAC	ø	CTCGGCGC	တ	CCGGTGGCT	O	GCT	H	555	ĸ
ڻ	3GT(>	GCA(I	000	Ω,	GGA	ជា	GAT	H
æ	ÖCC	R.	360	æ	252	4	909	æ	299	Ø
A I	3CG(K	3ATGGCGGC	æ	GTT +	لنا	CAC +	₽	CGA +	ជ
>))) (æ	SAT	Σ	GTG	3	CGA	. Ω	CGA	Ω
>	3GT(>	STC	တ	000	Æ	GGT	>	CTT	Ĺų
Ω	#CC(Q	AATGTGT +	>	CCT	1	7GC	Æ	900 -+-	&
Ω	CCGGAACCGGTCGCGCCCGGGTCCAGGAGCTCTCGG +++++	ធ	CAA	z	crcssccrsscsrssrcscsc	æ	CCTCACCGGAGGTGCGGTCGACACGCGGGAGCTGGTGGCCGACT	ტ	GGCTGCCGGTGCGCTTCGACGAGGCGATCCGGTCCGCCTGG	>
Ω	30CC	Δ	CAT	H	OH C	လ	990	Ŋ	000	Д
3	3GTCG	S	AGGT(-+	>	rgcg(%	CAC +	Ħ	GCT +	н
C.		ტ	ACA.	ø	GAT	Σ	CCT	႕	555	∝
æ	GAC	H	292	æ	999	Ŋ	CAG	တ	CTT 	Ĺ
н	GCT	H	900	ĸ	CGA 	ជ	292	Ø	CAG	လ
×	FCTC +-	Н	GGGGTCC	>	+	4	CTA +	×	CGCCGCAG	ሺ
Д	3G T	>	36G	ტ	CAT	н	CIT	بنا	 909	ĸ
ា	GCTCGGTTCTGCTGACCGGGTCGCCGGAACCGGTCGCGCGCCGGGTCCAGGAGCTCTCGG	တ	CCGAGGGGGTCCGCGCACAGGTCATCAATGTGTCGATGGCGGCGCGCACTCGGCGCGCAGGTCG	ធ	ACGACATCGCCGAGGGGATGCGCTCGGCCCTGGCGTGGTTCGCGCCCGGTGGTCGGTCGGAGG	Д	TGCCCTTCTACGCCAGCCTCACCGGAGGTGCGGTCGACACGCGGGGAGCTGGTGGCCGACT	а	ACTGGCGCCGCAGCTTCCGGCTGCCGGTGCGCTTCGACGAGGCGATCCGGTCCGCCCTGG	3
	1381		1441		1501		1561		1621	



			5	/89 -	7-7			٠	
2100	2160	1	2220	i i	F16.2-	ı	2340	1	2400
AAGCCGACGCCGAGGCCACGTTCCGCGAGCTGGGGCTGGACTCGGTGCTGGCCGCGCAGC+++++++	TGCGCGCCAAGGTGAGCGCCGCGATCGGGCGCGAGGTCAACATCGCCCTGCTCTACGACC	RAKVSAAIGREVNIALLYDH	ACCCGACTCCGCGTGCGCTCGCGGAAGCACTCGCGGCGGGAACCGAGGTCGCACAACGGG	PTPRALAEALAAGTEVAQRE	AAACCCGCGCGCGGACCAACGAAGCGGCGCCCGGCGAACCGGTCGGGGTCGTCGTCGCGATGG	TRARTNEAAPGEPVAVAMA	CCTGCCGGCTGCCCGGCGGTGTGAGCACCCCCGGAGGAGTTCTGGGAGCTGCTGTCGGAGG	CRLPGGVSTPEEFWELLSEG	GCCGCGACGCGGTCGCGGGACTGCCGACCGGCGGCTGGGACCTGGACTCGCTGTTCC
2041	2101		2161		2221		2281		2341

						2 - 6	ĺ				
ı	2460	1	2520	i	2580	- FIG 2-	2640	1	2700	ı	2760
RGWDLDSLF	CTCGGGCACCGCGCACCAGCGCGGCGGCGGTTTCCTGACCGAGG	Q R G G G F L T E	TCCCCGCGCGAGGCGCTGGCCGTCGACC	SPREALAVD	GAAGTGCTGGAACGGGCGGGAATCCCCGC	EVLERAGIP	STTCGTCGGCCTGATCCCGCAGGAGTACG	F V G L I P Q E Y	AGGGCGCGAAGGCGTCGAGGGCTACCTGATGACCGGTACGACCA	EGYLMTGTT	SCTCGGCCTGGAGGGCCCGGCGATCAGCG
RDAVAGLPTD	ACCCCGACCCCACGCG	PDPTRSGTAH	CGACCGCGTTCGACCCGGCCTTCTTCGGCATGTCCCCGCGCGCG	TAFDPAFFGM	CGCAGCAGCGGCTCATGCTCGAGCTCTCCTGGGAAGTGCTGGAACGGGCGGG	QQRLMLELSW	CGACCTCGTTGCAGGCCTCGCCCACTGGCGTGTTCGTCGGCCTGATCCCGCAGGAGTACG	T S L Q A S P T G V	GCCCGCGGCTGGCCG1	PRLAEGGEGV	CGAGCGTCGCCTCCGGCCGCATCGCCTACACGCTCGGCCTGGAGGGCCCGGCGATCAGCG
	2401		2461		2521				2641	•	2701

6/89 🕠



					8	/89	7				
							FIG.2-7				
3120		3180	ı	3240) 	3300	1	3360	1	3420	ı
+	a	55 +	O) 1	· ¤	991	æ	H H	Д	TCA 	လ
1	н	GCACG	Ħ	TACGGGC	Ŋ	CCAGG	ø	ည	н	GAJ	Н
1	>	360	Æ	GTA	×	CAC	E	CAC	.H	7GC	K
į	ĸ	CGA(回	360	. 4	CCA +	I	999 +	O	4TCCGG	G
i	>	CCGTCGA	>	CGACCCGATCGAGGCGCGCGCTGTTCGAGGCGTACGGGC	ធ	CGGCCA	ပ	TCAAGATGGTGCTGGCGATGCGCGCGGGGCACCCTTC	4	ATC	တ
	α	292	æ	GTT	ŢŦ	CCT	H	909	K	GTC	လ
	æ	CGA	Q	CGGGCGCT	ы	CAA	Z	GAT	Σ	CGACTG +	3
+	æ	CATCG	H	395 	A	GTCCA	S	999	K	CGA	Q
İ	ט	CGA	Q	3CG	ĸ	CAA	×	GCT	H	GAT	н
į	Z	292	æ	2000	Æ	GGT	>	GGT	>	GGA	田
+	Δι	92229995	A	TCGAGG	. ш	3GC1CG	လ	GAT -+-	Σ	GAA -+-	×
	Æ	CGG	ပ	GAT	н	999	U	CAA	×	GTC	တ
	တ	ICT	н	222	Δı	CCT	H	GAT 	н	909	ĸ
1	н	CGG	Ŋ	CGA		GCA +	Ħ	CGT +	>	GGA +	回
i I I	ڻ .	GTC	လ	CGG	O	GCT	н	990	ტ	ATC 	တ
) 	z	AGA 	回	ACT	H	225	а	TGC	K	292	K
1	တ	299	4	900	ĸ	GCA	a	TGT 	>	GCA	H
+	A.	GCTG	н	TACCC	H	CGA(ធា	+	Ŋ	TCT(+-	н
i 1	U	3000	K	366	O	500	ĸ	282	æ	CAC	H
 	Ω.	AGCAGGCGCTGGCAGAGTCCGGTCTCGGGCCCGCCGACATCGACGCCGTCGAGGCGCACG	a	GCACCGGTACCCGACTCGG	E	GCGACCGCGAGCAGCCGCTGCACCTGGGCTCGGTCAAGTCCAACCTCGGCCACACCCAGG	Ω	CGGCCGCCGGTGTTGCCGGCGTGATCAAGATGGTGCTGGCGATGCGCGCGC	A	CCCGCACTCTGCACGCATCGGAGCGGTCGAAGGAGGAGTCGACTGGTCATCCGGTGCGATCA	. ¤
3061		3121		3181	1) 1	3241		3301		3361	

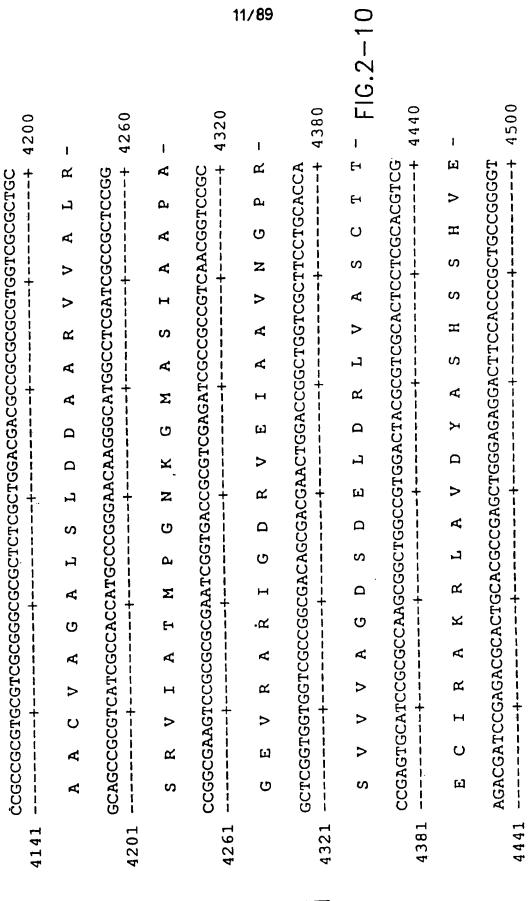
۲,

GGGTCT + 3480	- 8 >	AGGTCG	- ^ ^ Q	SCGAGCA	A S S - FIG 7 - 8	36	н Р G	GCCCCACC + 3720	P H R -	CTCGCGA + 3780	L A T -
GCCTGCTCGACGAGCCGGAGCCGTGGCCCGCCGCGCGCGC	R A G	CGTCGTTCGGCATCAGCGGCACCAACGCGCACGCCATCATCGAGGAAGCTCCGCAGGTCG	a a	TCGAAGGCGAGCGGGTCGAGGCCGGCGACGTCGTGGCGCCCCTGGGTGCTTTCGGCGAGCA	r S	GCGCGGAAGGTCTGCGCGCCCAGGCGCGCGCGCTGGCCGCGCACCTGCGCGAGCACCCCG	ය ය	TCGCGTACTCGCTCGCGACGGGACGGGCCGCGCTGCCCCACC	A L	TCGACGAGTCCGCGCGCTGCGCGTGCTCGACGGTCTCGCGA	L D G
GACCGC	P R	CATCGAGG.	교	CGCCCTGGG	P W V	CGCGCACC	A H L	GGGACGGG	G R	CGCGTGC	R V I
\$ 90909099	G A R	GCCATCA	A I I	GTGGCGC	V A E	נדפפככנ	L A 1	GCGACG	A T	GCGCTG	A L
3000000	K A	CGCGCACGC	H	CGACGTC	D Q	+	A R	CICGCIC	S	GTCCGCC	S
AGCCGTG(M d	GCACCAA(z	AGGCCGG	g K	CCCAGGCGGCGCGGCTG	Ø	TCGCGTA	A	TCGACGAG-+	ы О
SAGCCGG.	ਬ ਰ ਜ਼	ATCAGCG	I S G	CGGGTCG	я V Е	CTGCGCG	L R A	CGCGACA	R D I	CCTTCGCCCCCG	A P V
GCTCGAC	L D	GTTCGGCAT	F G	AGGCGAG	<u>ធ</u>	GCGCGGAAGGTCTGCGCG	ហ	GTCAGGACCCGCGCGACA	Б	9 - -	A FI
3421	н	CGTC 3481	v	TCGA 3541	ជ	3601		GTC?	Ø	GCGC 3721	A

					•		FIG 2-0	, 			
3840	1	3900	1	3960		4020	- FIC	4080	ı	4140	4
	>	0 †	ል	5 + 1	臼	6 + +	H	TCGA	Σ	FCG +	A
CIT	Ŀı	CICCC	လ	CTT	៤	GTC	S	GTC	S	\GA]	H
CGT	>	CAC	H	GGA	Ω	GCI	н	999;	Ø	3CG7	ជ
+	K	CGA +	Ω	TCT +	H	099	K	:GC1	S	9991	Ŋ
TGCCGCCGTTGGAACGAGCCGGGCGCAGCAGCGCGCGCGTCTTCG	æ	TTGGCAGTGGGCGGCCATGGCCGTCGACCTGCTCGACACCTCCC	ы	GTTGCGCGAGTGCGCCGACGCGCTCGAACCGCATCTGGACTTCG	Ħ	GCGCGCGGAAGCCGCGAGGCGGGAGCAGGACGCGGCGCCTGTCGA	æ	GGTGCAGCCCGTGATGTTCGCGGTCATGGTCTCGCTGGCGTCGA	ß	CGGGCACTCCCAGGGCGAGATCG	ø
SCA	a	CCT	ы	ACC	Д	GGA	Ω	GGT	>	CTC	လ
SCA	ø	CGA	Ω	CGCGCTCGA	ធ	GCA	ø	CAT	Σ	GCA	Ħ
GGCGCA	æ	CGTCGA	>	GCT +	Ä	GGA +	ធ	GGT +	>	+ CGG	ပ
) 	æ	299	A	1 1 1	K	909	ĸ	292	æ	CAT	н
GAG	တ	CAT	Σ	CGA	Ω	AGCCGCGAG	ĸ	GTT 	(tu	CGTCGAGCCGGCCGCGGTCAT	>
GGAACG	H	+	ပ	-+- -+-	K	- -	K	GAT	Σ	-+- CGC	A
rgg.	U	360	æ	GTG	U	AGC 	4	CGT	>	299	A
E CGT	>	37G	3	CGA	ក	GGA	ធា	225	Д	229	а
+ +	æ	3CA(a	GTTGCGCGAGTGCGCCG	æ	250505	A	GCA +	ø	CGA +-+	ជ
IGC	æ	ITG.	3	GHT	H	909	¤	GGT	>	CGT	>
	r	366	Ŋ	Ö U U	K	CCT	н	CGT	>	555	Ŋ
CGA(Ω	SCA(α	AGC	K	GTT	៤រ	GGA	Д	CCA	Ħ
CGCC	Æ	-+	ט	CGCA	K	5000	Д	CGTG +-	>	AGCCCA	Ø
AAA	Z,	Ö	Δ	rtt(Щ	SAT(н) 3 2 3 1	ĸ	300	ĸ
CGGGAAACGCCGACGG	ຶ	TCTTCCCCGGGCAGGG	ĺΨ	CGGTTTTCGCAGCCGC	>	AGGIGAICCCGIICCI	>	CCGAGCGCGTGGACGT	េ	TGTGGCGAGCCCACGGCGTCGAGCCGGCCGCGGTCATCGGGCACTCCCAGGGCGAGATCG	3
3781		3841		3901	·	3961		4021		4081	

10/89

δ



					127 00			•			
	0		0		0		0	(FIG.2-		0
1	4560	1	4620	ı	4680	1	4740	ı	F1G	. 1	4860
псвоенргрсг	TCGTGCCCTTCTTCTCCACCGTCACCGGGCGCTGGACGCAGCCGGACGAGCTCGACGCCG	GRWTQPDELDAG	GGTACTGGTACCGGAACCTGCGCCGCACCGTGCGGTTCGCGGACGCCGTCCGT	R T V R F A D A V R A L A-	CCGAGCAGGGATATCGCACGTTCCTGGAGGTCAGCGCGCACCCGATCCTCACCGCCGCGA	LEVSAHPILTAAI-	SGGATCGGGCGCCGACCTCTCCGCCATCCATTCGCTGCGCCGCG	G A D L S A I H S L R R G -	GTGACGGCAGCCTCGCGGACTTCGGCGAAGCGCTCTCCCGCGCGTTCGCCGCCGCGGTGTCG	GEALSRAFAAGVA-	CGGTGGACTGGGAGTCGGTGCACCTGGGCACCGGAGCACGCCGGGTGCCCTTGCCCACCT
A	TCA	T	0000	K	ricc	FI I	rcgo	S	rrce	Fi.	CAC
LH	ACCO	T	CIGC	I I	ACG1	H	GGA.	ს	GAC:	Ω	GTG(+
	TCC.	S	AAC(z	292	~		Ω	909 1	Æ	TCG
Ω	TIC	Ĺц	550	K	TAT	×	2993	Ŋ	CTC	H	GAG
~	TTC-+-	ĮΞι	TAC	×	GCAGGGATA +	Ŋ	SATC	н	SAGC	တ	7TG
н	3000	Д	CTGG	B	3CAG	α	3GAC	ធា	0000	ប	3GA(
H	TCGTC	>	GGTAC	×	CCGAC	ធា	TCGAGGAGATCGGCGA	ជា	GTGACGGCAGCC	Ω	CGGT(
	4501		4561		4621		4681		4741		4801

12/89

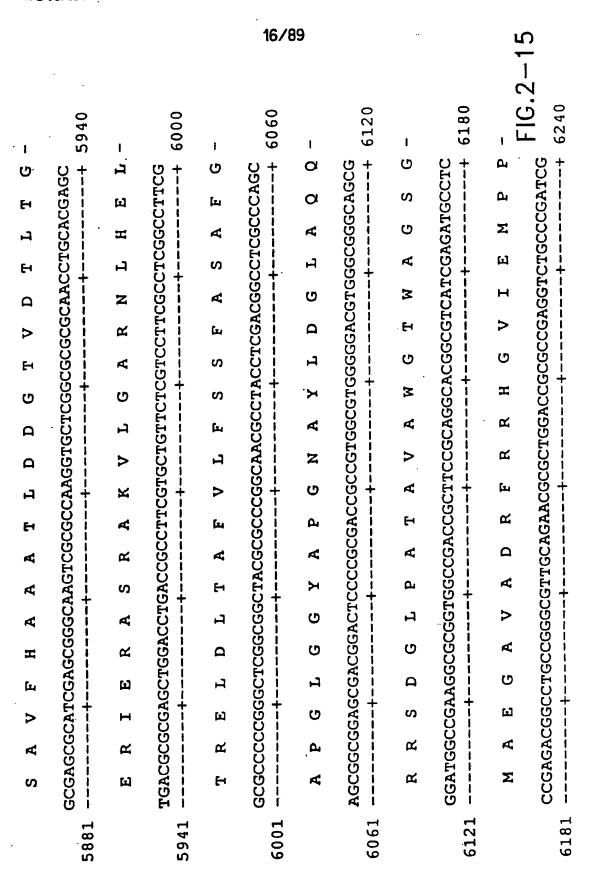
					13/89			,	7.1-	
ı	4920	ı	4980	1	5040	ı	5100	i i	FIG.2-12 5160	
×	ACCG +	ы	7G -+	回	GA ++	H	50 +	>	9 1 1	A
F	CAC	H		O	CGA	ជ	GGT	>	GGT	>
ы	3TC(ဟ	IGC.	æ	3GA	Q	ACT	ы	CGA 	េ
H) 	ĸ	1992 1+	U) 1 1 1 1	Ø	CGA +	បា	+	Ŋ
Д	3000	æ	TCGAGTGGCGGCCCACCGGTGCCGGTG	H	AAC	H	AGGCCCTGGAGTCGGCCGGGGCGCGGGTCCGCGAACTGGTCG	ĸ	GCGACGAACTCGCGGAGCGGCTTCGTTCGGTCGGCGAGGTGG -++++	>
>) 1 1 1 1	Ø		בי		ប	CGT	>	TTC 	လ
R	GCCGGTG +	>	 929	æ	292	4	909	ĸ	AGCGGCTTCGT	K
æ) +	Д	GTG 	3	GTA +	X	+ 299	K	GCT +	J
æ	GAA)	×	CGA	ធ	GAA	×	993	ß	909	ಜ
v	SCGTCTGGCTCGAACC -++-	Д	CAT	н	299	K	299	Æ	GGA	<u>ធ</u>
H	CGA -+-	ഥ	CTACCGCA'	æ	GGT -+-	r . v	GTC -+-	လ	ACTCGCGG	æ
Ö	GCT	H	CIA	×	GCT	ı	GGA	ы	ACT	Н
H	CTG	3	 929	æ	CIG	3	CCT	ı	CGA	ы
Ħ	CGT +	>		ы	CAC	H	+: 2991	K	3CGA -+	Ω
>	909	K	00	Æ	990	Ŋ	GGA	ធ	HCG -	æ
ഗ	CGA	凹	TTC	ဟ	CGA	Q	TCG	æ	990	Ŋ
ы	TTCCAGCG	¤	GACGAGGITIC	>	GCT	1	299:	æ	GCCCGCTG +	ပ
X	CCA	ø	CGA	回	900	ĸ	292	A	9 - - -	ø.
Ω.	GTT	្រា	CGA	Ω	292	K	CAC	H	292	æ
>	ACCCGTTCCAGCGCGAGCGCGTCTGGCTCGAACCGAAGCCGGTGGCGCGCGC	പ	AGGTCGACGAGGTTTCCGCGCTGCGCTACCGCATCGAGTGGCGGCCCACCGGTGCCGGTG ++++++	>	AACCCGCCCGGCTCGACGGCACCTGGCTGGTGGCGAAGTACGCCGGAACCGGGACGAGA	Сı	CGAGCACCGCGCTCGGGAGGCCCTGGAGTCGGCCGGGGGCGCGGGTCCGCGAACTGGTCG	ഗ	TGGACGCCCGCTGCGGTC	Ω
	4861		4921		4981		5041		5101	

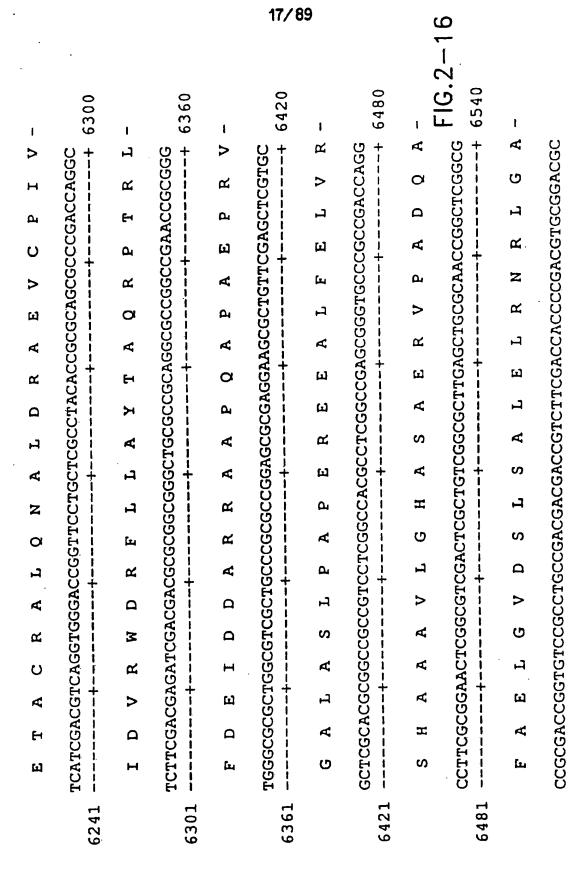
3

	CGGCGGTCGTGTCCGGCGCCGCCGGTGAGGACCAGCTCGCGCTGCGCGCCGACGGGGTGT F16.2-1+++++++++++++++++++
CCG	CGGGGTGT +
	CGGGG1
	5 6
L L L CGG CGG A H H	
SCGCCC SCCGA SAACG SAACG SAACG C-+ B R B R	CCGA -+
E A E TCGGCCGA TCGGCCGA TTCGAACG TTCGAACG TTCGAACG TTCGAGGAA T E R T E R T E R T E R T E R T E R T E R T E R T E R T E R T E R T E R	CGC
S S S S S S S S S S S S S S S S S S S	30.00 R
GGGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	3CT(
SAGCCGGAGGGCGGAGGGCGATGGTGGGGGCCGGGGCCGGGGCCGGGCCGGGCCGGGCCGGCCGGGCCGGGCCGGGCCGGGCCGGGCCGGGCCGGGG	CGCGCTG
D E A E P E E A P D CCTCGTGCAGGCGAGGCCGGAAG CCTCGTGCAGGCGATGGTGTCGGCCGAAG L V Q A M V S A E D A V A T G P F E R R CGTCGGCGGCGTCATCGCGCTGGAGAC CGTCGGGCGGTCATCGCGCTGGAGAAC CGTCGGGCGGTCATCGCGCTGGAGAAC A V A T G P F E R CGTCGGGCGGTCATCGCGCTGGAGAAC CGTCGGGCGGTCATCGCGCTGGAGAAC CGCGGGGTCGTCGCCGGGCTCGGCCCGG A G R V I A L E N CGCGGGGTCGGTCGCCCGG A G S V A E L A R	CAGCTO
STGG STGG STGG STGG STGG STGG STGG STGG	CCA
ACGAAGC LE A TCGTGCA V Q CCGTCGC L+ G R CGGGGTC CGGGGTC G S CGGGGGTC	GGA -+- D
D E COTCG C C C C C C C C C C C C C C C C C C	TGA E
S S S S S S S S S S S S S S S S S S S	555
26CC 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	90000 -+
T T T T T T T T T T T T T T T T T T T	9000
1 L L L L L L L L L L L L L L L L L L L	990
3 S S S S S S S S S S S S S S S S S S S	STC
GGAGTGCTGTC G V L S GCTTCGCTGGC GCTTCGCTGGC A S L A CCGCTGTGGAC CCGCTGTGGAC TGGGCGCCCT TGGGGCGGCCT W G G L	7 7
2 S S S S S S S S S	3676
G V L S L L A V D E A E P E E A P L A L TGGCTTCGCTGGCGCGCCCTCAGCCTCGTGCGCGCGAGGCGCCGCTCGCGC TGGCTTCGCTGGCGGACACGCTCAGCCTCGTGCAGCGCGAACTCGGAT A S L A D T L S L V Q A M V S A E L G C GTCCGCTGTGGACGGAAAGCGCCGTCGCGACGGGGCCGTTCGAACGCGTCGCA P L W T V T E S A V A T G P F E N P N ACCCCGCCCACGCCCTGTGGGGCGTCGGCGGGCGTTCGAAACCCCGCG A A H G A L W G V G R V I A L E N P A V TGTGGGCGCCTGGCGCGCGCGCGGGGGCTCGCGGACCCGGCCCTCGGAAACCCCGCG W G G L V D V P A G S V A E L A R H L A	CGGCGGTCGTGTC
5161 5221 5281 5341 5401	5461

14/89

				15/89					F IG.2-14 5880
5580	5640)))	5700	1	5760	ı	5820	יו נ	
ACGGACGCCGGTGCGCGCGCGCGCCCCGGCGACCGATGAC	G R R W V R A A P A T D D E W R F I GAACCGTGCTGGCTGCTGGCGGCGGGGGGGGGGGGGGGG	TVLVTGGTGGVGGQ	CCCGGCGGGGCGCCCCCACCTGCTGGTGAGCCGCAGCGGGCCGGACGCGGACGCGGCGGTAL	RRGAPHLLLVSRSGPDADGA	CCGGCGAACTGGTCGCCGAGCTCGAGGCGCTGGGCGCCCGGACGACCGTCGCGGCCTGCG	GELVAELEA.LGARTTVAACD	ACGIGACCGACCGCGAGICGGITCGCGAGCIGCICGGCGGCAICGGIGACGACGICCCGC	V T D R E S V R E L L G G I G D D V P	TCTCGGCGGTGTTCCACGCCGCCGCCACGCTCGACGACGGCACCGTGGACACCCTTCACCG
5521		5581	5641		5701		5761		5821



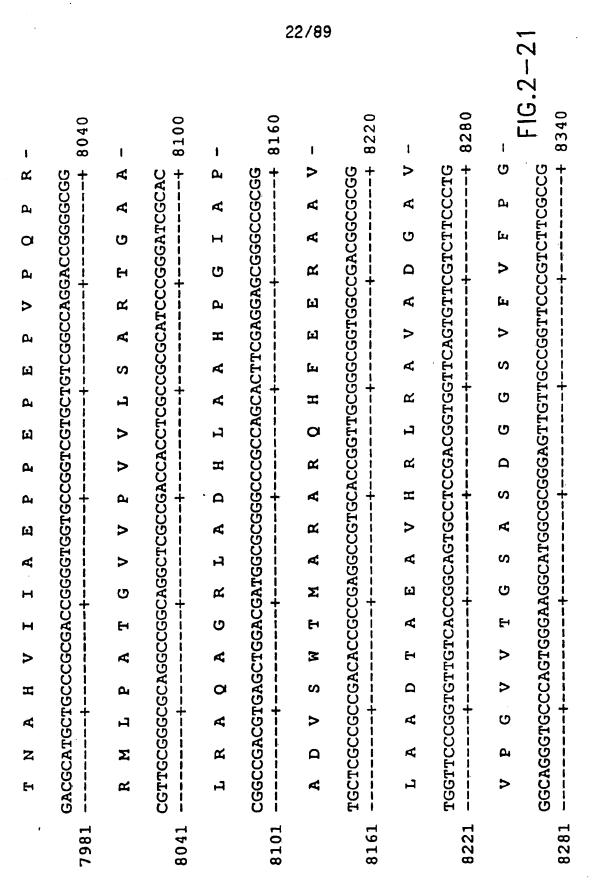


					18/	89			!	-17	
0099	1	0999	1	6720	1	6780))	6840	, i	FIG.2-17	
+	н	993	4	+ 5000	G	590	• 4	+ 99299	ტ	TTCG	· · · · ·
	Ħ	AGCAGGCGGCACCGG	д	1GCC	ρı	CACCTCGGGACGCGACTCCGCGG	K	 	K	ICT	
	ĸ	3660	Ø	ည်ဗ	H	ACTO	ဟ	ACG(æ	CGCGGCGTT	រុំក
+	>	999+	A	GTGCCG +	K	15 CG	Ω	TCCGA	Ω))) () ()	K
	Ω	CS	a	GTG	ပ	ACG	ĸ	CTC	လ	000	K
1	Δı	CGA	ы	099	4	999	O	999	A.	250	Ω
	H	ccgaactcggcggtgcgaccggagccg	4	CAT	Σ	CIC	တ	AGCTGAT	Σ	CTI	Ĺτι
+	Q	+ 550	Ŋ	CGG(O	CAC	H	GCT	H	TGAC	Q S
1	Įщ	GAC	H	CGT	>	GAT	н	CGA	ធា	990	ပ
1 1	>	TGC	4	GAT	н	TGTGGGAGCT	Ä	CGA	Ω	090	K
+-	H	-+-	ပ	- + -	Æ	GGA -+-	. ш	2222			U
!	H	990	Ŋ	GAT	H	GTG	×	GGT	>	099	K
) 	H	ACT	н	229	α	GCT	J	CTG	3	CAT	Σ
<u> </u>	.Δι	CGA +	ធ	CGA +	ធ	AGCGG	<u>~</u>	3666		CTT +	[LI
i i	н	000	4	CGA	Ω	GGA	ធ	900	ĸ	CAA	Z
 	¤	099	æ	CGT	>	222	а	TGA	Ω	990	ဗ
1	>	CCT	H	O C C C C	Δ	CTC	တ	CGA	Ω	CCA	H
+	Ŋ	GCA +	Ħ	20000	A	GGTCGACT	Ω	-+	Дı	CGCC	Æ
İ	H	395	4	CAC	EH	GGT	>	EGT	>	CCG	ជ
1	æ	TGGCGGCGCACCTGGCCGCCGAACTCGGCGGTGCGACCGGAGCCGAGCAGGCGGCGCACCGG	4	CGACCACGGCCCCCGTCGACGAGCCGATCGCGATCGTCGGCATGGCGTGCCGGCTGCCCG	H	GGGAGGTCGACTCCCCGGAGCGGCTGTGGGAGCTGATCACCTCGGGACGCGACTCCGCGG	ធ	CGGAGGTCCCCGATGACCGGGGCTGGGTCCCCGACGAGCTGATGGCCTCCGACGCGGGGGGG	<u>면</u> .	GAACCCGCGCCCACGGCAACTICATGGCGGGCGCCGGTGACTTCGACGCGGCGTTCTTCG	E→
6541		6601		6661		6721		6781		6841	

					19/	89				FIG.2-18	
0969	1.	7020	1	7080	ı	7140	ı	7200		FIG.	1
GGATCTCGCCGCGCGCGCGCTGGCGATGGACCCGCAGCGCCCAGGCGCGCTGGAGACGA	ISPREALAMOPQORQALETT	CGTGGGAGGCGCTGGAAAGCGCGGGCATCCCACCGGAGACGTTGCGCGGCAGCGACACCG	WEALESAGIPPETLRGSDTG	GCGTGTTCGTCGGCATGTCCCACCAGGGCTACGCGACCGGGCGTCCGCGCCCCGGAGGACG	V F V G M S H Q G Y A T G R P R P E D G	GCGTCGACGGGTACCTGCTCACCGGCAACACCGCGAGCGTCGCGTCGGGACGCATCGCCT	V D G Y L L T G N T A S V A S G R I A Y	ACGTGCTGGGGCTGGAAGGTCCCGCGCTGACGGTGGACACGGCGTGTTCGTCGTCGTCGTTGG 	V L G L E G P A L T V D T A C S S S L V	TGGCGTTGCACACGGCGTGTGGGTCGTTGCGTGACGGTGACTGCGGTCTTGCGGTGGTGCCG	ALHTACGSLRDGDCGLAVAG
6901		6961		7021		7081		7141		7201	

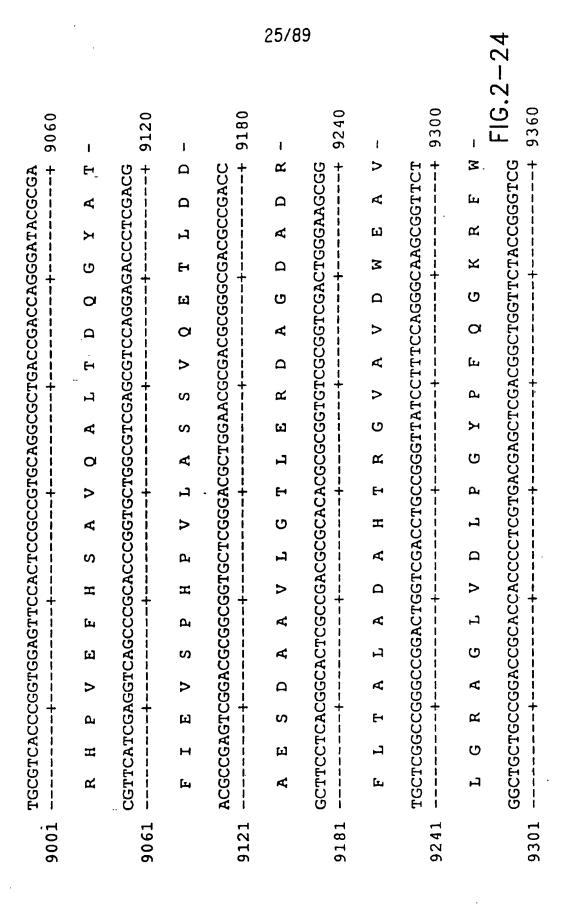
•					20/	′ 89			7	<u>1</u>	
7320	1	7380	1	7440	1	7500	1	7560	((FIG.2-19	į
GTGGTGTGTCGGTGATGGCGGGTCCGGAGGTGTTCACCGAGTTCTCCCGCCAGGGCGCGCGC	V S V M A G P E V F T E F S R Q G A L	TCTCGCCGGACGGCCGGTGCAAGCCCTTCTCGGACGAGGCCGACGGATTCGGTCTCGGGG	PDGRCKPFSDEADGFGLGE	AGGGTTCGGCGTTCGTCGTCCTCCAGCGGTTGTCCGACGCCAGGCGGGAGGGCCGCCGCCGCCGCGCGCG	SAFVULQRLSDARREGRRV	TGCTCGGCGTGGTGGCCGGGTCCGCGGTGAACCAGGACGGCGCGCGAGCAACGGGCTCTCCG	G V V A G S A V N Q D G A S N G L S A	CTCCGAGCGGCGTCGCGCAGCGGGTCATCCGCCGGGCGTGGGCGCGTGCGGGGATCA	SGVAQQRVIRRAWARAGIT	CGGGCGCGGATGTGGCCGTGGTGGAGGCGCATGGGACCGGTACGCGGCTGGGCGATCCGG	A D V A V E A H G T G T R L G D P V
GTG 7261	o	TCI 7321	S	AGG 7381	U	TGCTC 7441		CIC 7501	<u>α</u>	CGG 7561	g

					21	/89			. (FIG.2-20
7680	1	7740	1	7800	ı	7860	i	7920	i	FIG.
TGGAGGCGTCGGCGTTGCTGGCTACTTACGGCAAGTCGCGGGGCGGTCGTCGGGCCCGGTGC	EASALLATYGKSRGSSGPVL	TGCTGGGTTCGGTGAAGTCGAACATCGGTCACGCGCGCCGCGCGCCGCGGGTGTCGCGGGCG	LGSVKSNIGHAQAAAGVAGV	TGATCAAGGTGCTGCTCGGCCTGGAACGCGGTGTGGTGCCCCCCGATGCTGTGCCGGGGCG	IKVLLGLERGVVPPMLCRGE	AGAGGTCGGGCCTCATCGACTGGTCCTCCGGCGAGATCGAGCTCGCAGACGCGCGTGCGGG	R S G L I D W S S G E I E L A D G V R E	AGTGGTCGCCCGCCGCGGACGGGGTGCGCCCGGGCAGGTGTGTCGGCGTTCGGGGTGTGAGCG	WSPADGVRRAGVSAFGVSG	GGACGAACGCGCACGTGATCATCGCCGAGCCGCGGAACCGGAGCCCGTGCCGCAACCGC
7621		7681		7741		7801		7861		7921



						22				
1	8400	I	8460	1	8520	ı	8580	ı	FIG.2-22	ı
· ы		>		ы	55 +	æ); 	7) 	æ
Æ	CGAGG	ធា	3GTGC	>	GCTGGCGCGTTGTGGCGTGCCTGCGGTGCCGTTCCTTCGG +++++++	တ	GGGTGAGATCGCCGCCGCGGTGGTGGCGGGAGCGTTGTCGC	S	CGTCGCCCGCCGGTCGAGGGCGGTGCGTGCGGTCGCGGGCCC+++++++	g
(TI	STC(ဟ))))	Д	TCC	а	GTT	ᆸ	292	A
>	3GT(>	GCA +	Ø	CGT +	>	AGC +	æ	.GGT	>
Д	CTCGG	တ	GTGGTGCAG	>	TGC	æ	9993	ပ	TGC	Æ
>	ATT	ក្រ	CGT	>	990	Ŋ	3660	A	2001	æ
	 990	Ŋ	CGGGTCGA	Ω	GCGTGCCTG	O	1991	>	5663	>
H	-+	æ	GGT 	>	TGC	æ	1991	>	9666	A
H	GGT	>	505	ĸ	525	~	2000	A	GAG	ĸ
ា	GGA	ធ	GGA	ធ	CGGTTGTG	3	200	Æ	CGCCGCCGGTC	S
ĸ	rgrcgg +	ဟ	CGCT	1	GTT	H	500	A)) (~
Ø	GIT	H	GTC	S	၁၁၅	æ	AGAT	H		ĸ
Σ	GGT	>	000	<u>a</u>	299.	A	TG	ជ	500	A
O	TGC +	Æ	292	K	GCTG	1	19991	Ŋ	CGT-+-	>
臼	CGA	Q	GGA	Ω	GIC	လ	(C)	α	30.67	>
3	GTG	Ö	TCC	а	GGT	>	CTC	လ	rGCC	ĸ
a	CGA	ច្ន	ACG	ĸ	GAT	Σ	AGGGC?	Ħ	CGGCA1	Σ
¥	-+	æ	000	а	199	>	AGG	G	0901	ß
Ŋ	GAT	н	GGA	ធ	292	æ	CAT	H >	GGA	Ω
or 	AGTCGATCGCCGAGTGCGATGCGGTGTTGTCGGAGGTGGCCGGATTCTCGGTGTCCGAGG	တ	TGCTGGAGCCACGTCCGGACGCCGTCGCTGGAGCGGGTCGACGTGGTGCAGCCGGTGC	ы	TGTTCGCGGTGATGGTGTC	្រ	CCGTCATAGGGCACTCGCA	. >	TGGAGGACGGCATGCGCGT	្ន
	8341		8401		8461		. 1028	ł 1)	8581	•

	-				24/	89			1	FIG.2-23	
8700	ı	8760	t	8820	ı	8880	1	8940	อ , i	F1G.2	1
GGGGGAGCATGCTCTCGGTGCGCGGCGGCCGCTCCGACGTCGAGAAGCTGCTCGCCGACGACG	G S M L S V R G G R S D V E K L L A D D	ACAGCTGGACCGGCAGGCTGGAGGTCGCCGCGGTCAACGGCCCCGACGCCGTGGTGGTGG	SWTGRLEVAAVNGPDAVVA	CCGGTGACGCCCAGGCGCGCGCGCGAGTTCCTGGAGTACTGCGAGGGCGTGGGCATCCGCG	G D'A Q A A R E F L E Y C E G V G I R A	CCCGCGCGATCCCGGTGGACTACGCCTCGCACACGCGCGCACGTCGAGCCCGTGCGCGACG	RAIPVDYASHTAHVEPVRDE	AACTGGTCCAGGCGCTGGCCGGGATCACCCCGCGACGGGCCGAGGTGCCGTTCTTCTCCA	L V Q A L A G I T P R R A E V P F F S T	CCCTGACCGGCGACTTCCTCGACGGCACCGAGCTGGACGCGGGCTACTGGTACCGCAACC	LTGDFLDGTELDAGYWYRNL
8641		8701		8761		8821		8881		8941	



26/89 -250996 9540 0096 9420 9480 1 ı 1 K ū AACTCGACGCCGAGGGCATCGACGCGCCACTGTGGACGGTCACCTTCGGCGCGCGGTCGACG CGGGCAGTCCGGTGGCCCCGCCCGGACCAGGCGAAGCTGTGGGGGGCTGGGCCAGGTCGCGT CCGGCGCCGAACCGGAGGTCACGCGCGCGTCGGCGGGCTGGTCGGTGACTGCGCGGGCG **rggigicgitgcicgccitgagggggaiggigcggigcaaacccitgigciggiggg** Ω **ACTGGACCGAGGTGCCGCGCTCCGAACCTGCCGCGCTGCGCGCCGTTGGCTCGTGGTGG** TGCCCGAGGGGCACGAGGAGGACGGCTGGACCGTCGAGGTGCGGTCCGCGCTCGCCGAGG K Ω 回 > > K K K Н O П × G Ø > Ω 3 [s4 G Н Ŀ S 3 K Н H ፈ C G Ø > K > Н Ω ы G > Ħ Ч H ß 4 3 > 回 Ø C Н K H > a Ö Ω Д 3 K Ы G æ ы U 民 а Ω ы Ω H S H H H K ω > Н 回 K C K 回 Д Ы I ы Δ Д G Ы A 团 Д Н ഗ Ω Ы Н 9541 9361

27/89 FIG.2 - 260966 9780 9840 0066 ı 1 K ACACCGAGGGCGTCGGCGACCTGACCGCCGAGCTGACCCGGCTCGGCGCGCGGGTGTCGG CCCGGTGGCTCGCCGGTCGCGGCGCCGAACACCTCGCGCTGGTCAGCAGGCGAGGCCCGG CGGTGCCGGGCGGCACAATCCTGGTCACCGGTGGCACCGCCGGCCTGGGCGCGCGGAGGTGG Ŀ COGACGCCGTGCGTGCGCGGCGCTTTCCCCCGCCCACGTCACCGCCACCTCGGAGTACG ഗ AACTGCGAGGCCGTCTCACCGCGGTGCTGGCCGGCTCGGAGGACCAGGTCGCGGTGCGCG CCCTGGAACGCGGCCCCGCTGGACCGGCCTCGTCGACCTGCCGCACATGCCGGACCCGG ۵ > K Д G ы Ш Ω Ø K S O Д U K Σ > H G П ഗ Ø K Ξ G Ω Н G ۵ Ø 3 Ш F K S 二 Ч G G A × G Ξ J K K Ē Н O G Ч ഗ > > K Н Д Н G K 3 R Д K 瓦 K Ħ K Н H G Ы K K Д G K X G K G Н K G Д 3 S ы K K G H Ω 9901 9781 9841

					2	28/89			1	/7	
10080	1	10140	1	10200		10260	Ī	10320	C L	FIG. 2-2/ 10380	1
++	EGVGDLTAELTRLGARVSV.	CGTGCGACGTCAGCCGCGAACCGGTGAGGGAACTCGTGCACGGCCTGATCG	A C D V S S R E P V R E L V H G L I E	AGCAGGGCGACGTCGTCGGGGTGTGCACGCGGCGGGACTGCCGCAGCAGGTCGCGA	G D V V R G V V H A A G L P Q Q V A I	ACATGGACGAGGCCGCCTTCGACGAGGTGGTCGCGGGCCAAGGCCGGGGGGCGCGG	D M D E A A F D E V V A A K A G G A V	TGCACCTGGACGAGCTGTGCTCGGACGCCGAGCTGTTCCTGCTGTTCTCCTCGGGGCCG	LDELCSDAELFLLFSSGAG	GGGTGTGGGGAAGCGCCCGCCAGGGCGCCTACGCCGCGGGCAACGCGTTCCTGGACGCCT	WGSARQGAYAAGNAFLDAF
; { ! ! !	H	TGCACG	Ħ	AGCAG	a	TCAACG	Z	TGCAC	Ħ	GGGTG	>
10021		10081		10141		10201	•	10261		10321	

			2	9/89			•	-28	
10440	10500	- 10560	1	10620	t	10680	1	FIG.2-28	1
TCGCCCGGCACCGCGGGCCGCGCCTGCCCGCCACGTCGGTGGGGGGCTGTGGGGGCTGTGGGGGCTGTGGGGGCTGTGGGGGCTGTGGGGGCTGTGGGGGCTGTGGGGGCTGTGGGGGCTGTGGGGGCTGTGGGGGG	cgaggaggccgrgrcgrrccrgcgcgagcgcggrgrg ++	D E E A V S F L R E R G V R GCGCCCTCGCCCTGGACAGGGTGCTGGCCTCCGGGGAGA -+++++	A L A A L D R V L A S G E T	CGGCGGTGGTCGTGACGGACGTGGACTGGCCCGCCTTCGCCGAGTCCTACACCGCCGCCC	V D W P A F A E S Y T A A R	GGCCCCGGCCGTTGCTCGACCGCATCGTCACGACCGCGCGCG	RIVTTAPSERAGEP	CGGAGACGGAGAGCCTGCGCGACCGGCTGGCGGGTCTGCCCGCGTGCCGAGCGGACGGCGG	D R L A G L P R A E R T A E
		GGGCGATGCCCGTACCGC	AMPVPRA		AVVVTDV		PRPLLDR	CGGAGACGGAGAGCCTGC	ETESLRD
10381	10441	10501		10561		10621		10681	

					3	0/89			(FIG.2-29	
10800	ī	10860	1	10920	1	10980	ı	11040	i		ا ص
AGCTGGTGCGCCTGGTCCGCACCAGCACCGCGACCGTGCTGGGCCACGACGACCGAAGG	A T V L G H D D P K	CGGTGCGCGCGACCACGCCGTTCAAGGAGCTCGGGTTCGACTCGCTGGCGGCCGTCCGGC	LGFDSLAAVR	TGCGCAACCTGCTCAACGCGGCCACCGGGCTCCGCCTGCCGTCGACGCTGGTCTTCGACC	LRLPSTLVFD	ACCCGAACGCCTCCGCGGTCGCCGGTTTCCTCGACGCCGAGCTCGGCACCGAGGTCCGGG	LDAELGTEVR	GGGAGGCGCCCTCGCCCGGGCTGGACGCGCTGGAAGGCGCCCTGCCCGAGGTGC	DALEGALPEV	CCGCAACCGAGCGGGAAGAGCTGGTACAGCGCTTGGAACGGATGCTCGCCGCGCTACGCC	RLERMLAALR
	LVRLVRTST		VRATTPFKE		RNLLNAATG		PNASAVAGF		E A P S A L A G L		ATEREELVQ
10741		10801		10861		10921		10981		11041	

		31/89	-30	
iG ++ 11160 G -	11219	3 - 12702 -	c 12762 - 12762 - FIG.2-30 - 12822	ı ڻ
CGGTCGCCCAGGCCGCCGCCTCCGGGACCGGCGCCCAACCCGTCCGGCGACGACCTGG+++++++	ACGAACTGCTCGAAGCACTCGACGGCGATCGACGGCGATGCTCGACGCCGATGCATCGACTCGACGCCGGAGCTCGACGCGCGATGCATCGACGCGCGATCGACGCGCGATCGACGCGCGATCGACGCGCGATCGACGCGCGATCGACGCGATCGACGATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCATCGACGAATGCAATGAATG	CCGCCGATTGGAGAAAGGTGACTGACAGCGAGAAGGTGGCGGAGTACCTCCGTCGGGCG	ACGCTCGACCTGCGTGCCGCCGCGCGTCGCGAGCTGGAATCCGACCCGATCGCC+++++++-	I V S M A C R L P G G V N T P Q R L W E CTGCTGCGCAGGGCGTGAGACGCTGTCGGGCTTCCCCACCGACCG
CGGTCGCCCAGG	GCGAGGCGGCG'	CCGCCGATTGGA	ACGCTCGACCTG+ T L D L ATCGTCAGCATG	I V S M CTGCTGCGCGAG
11101	11161	12643	12703	

5

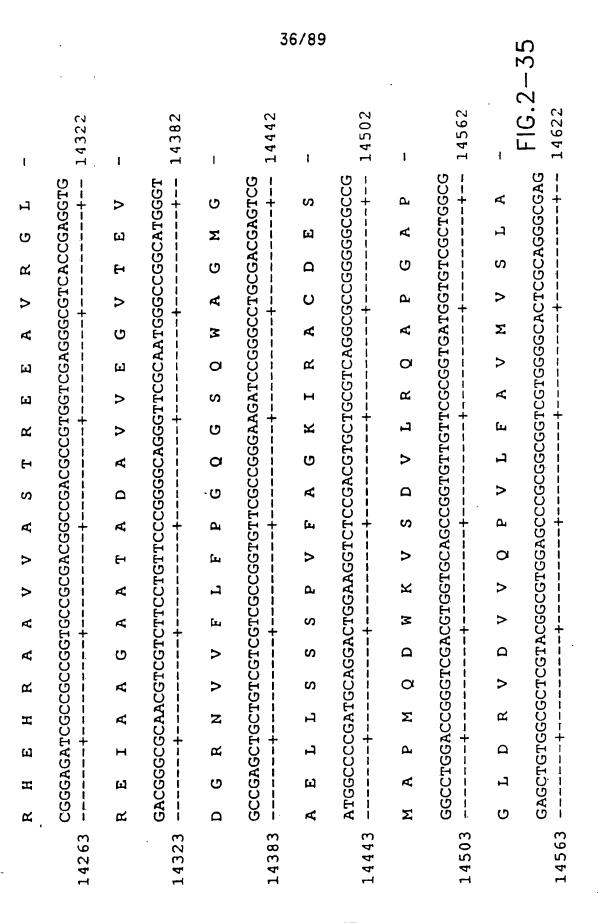
13062 12942 13002 i ı AAGTICGCCTACGGCGAGGACACCGCCGCGGGGGGGGCGTCGAGGGCTACTCGGTCACC GCCGGCATCGACCCGCACTCGCTGCGCGGTACCGCGACCGGCGTCTTCCTCGGAGTGGCG GCGCGGCTGCACCACCCGGACCCGGACAACCCCGGTACCAGCTACGTCGACAAGGGCGGG GCGGCCATGGACCCGCAGCAGCGGCTGCTGCTGGAGACGAGCTGGGAGCTGGTGGAGAAC 1++ Z G > Ω U ы (L) G 노 S 3 K П ⊁ Ω Δ Н G G K > ഗ 回 يتا > ω 3 \succ S G S ß > H Ω لتا H H E Ωi Ø G ы G ш Ĺų ы H K S 4 S K Z K S Н Ω Ω H 又 K K Ы Ы Ĺų H H Q Ы Ω G S Ω 4 O I ω G A K G S I Д Ы I Ω Д Ω \succ ы Σ Q Н Ω K K Н K G Н بئآ K 4 又 12883 12943 13003 13063

32/89

				3	3/89			7.0	70.	
13242 -	13302	ì	13362	ì	13422	1	13482	, <u>,</u>	FIG.Z-3Z 13542	ı
GGTGTGGCGCCCCGCGCTCCCTACACCATGGGCCTGGAGGGCCG++++++	TCGATCAGCGTCGACACCGCGTGCTCGTCGTCGCTGGCGCTGCTGCACCTGGCGGTCGAG	SISVDTACSSLVALHLAVE	TCGCTGCGCAAGGGCGAGTCGTCGATGGCGGTCGTCGGCGGTGCCGCGGTGATGGCGACC	SLRKGESSMAVVGGAAVMAT	CCGGGGGTGTTCGTCGACTTCAGCCGGCAGCGCGCGCTCGCCGCCGACGGGCGGTCGAAG	P G V F V D F S R Q R A L A A D G R S K	GCGTTCGGTGCCGGCGCCGACGGGTTCGGCTTCTCCGAAGGCGTCACCCTGGTCCTGCTC	A F G A G A D G F G F S E G V T L V L L	GAGCGGCTGTCGGAGGCGCGGCGAAACGGGCACGAGGTGCTGGCGGTGGTTCGCGGCTCG	E R L S E A R R N G H E V L A V V R G S
13183 -	13243 -	.	13303 -		13363	-	13423	•	13483	~

13842 13902 13782 13662 13602 1 1 1 CGCAACGGGGAACTGCCCGCGACCCTGCACGTCGAGGAGCCCACGCCGCACGTCGACTGG ATCGGCCACACCCAGGCGGCAGCGGCGTCACCGGCCTGCTGAAGGTGGTCCTGGCGCTTG ACCTACGGCCGCGACCGCCGACCGGCCGCTCTGGCTGGGCTCGGTGAAGTCCAAC GAGGCGCACGGCACCGGTACGGCGCTCGGCGACCCGATCGAGGCGAACGCGCTGCTGGAC **AGGGTCATCCGGCAGGCCCTCGAGAGCTGCGGTCTGGAGCCCCGGCGACGTCGACGCGGTG** 3 Z K Ω S Q 又 K Ω エ > Þ Д S 4 Z Ω S G 노 S G H Д Ч Д ы Δ 3 ω Ы d G 回 J H Д S > Δ а \vdash G Н Ξ K O G C G Н Ω S Z K H Ø K 回 S K Ø Ω H Ø Н 4 G ĸ Д U ø Ω Ø Н Ω O ы G 民 H Ø ፈ G Ξ I C Z G Z K Н K 回 4 13723 13603 13663 13543

					35	/89				-54
13962	ľ	14022	1	14082	i	14142	1	14202	1 I	FIG.2-54
TCGTCCGGCGCGTGCCGCTGCTGGCGGGCAACCAGCCGTGGCGGCGCGCGC	S S G G V A L L A G N Q P W R R G E R T	CGGCGCCCCCGTGTTTCCGCGTTCGGGATCAGCGGGACGAATGCGCACGTGATCGTCGTCAG	RRARVSAFGISGTNAHVIVE	GAAGCTCCTGAGCGCGAGCACCGGGAGACCACCGCGCACGACGGCCGACCGGTTCCGCTG	EAPERETTAHDGRPVPL	GTGGTGTCCGCGCGCACGACGCGGCGGTTGCGGGCGCAGGCCGCCCAGATCGCCGAGCTG	V V S A R T T A A L R A Q A A Q I A E L	CTCGAACGCCCGGACGCCGCCGCCGGGGTCGGGCTGGGCCTGGCCACGACCCGCGCCCCGCCCCGCCCCCGCCCCCGCCCCCGCCCCCC	LERPDADLAGIVGLGLATTRA	CGCCACGAGCACCGCGCCGCGTGGTGGCATCGACCCGCGAGGAAGCGGTGCGCGGACTG)3+++++++
13903		13963		14023		14083		14143		14203



					37/	'89			-36	
1	14682	1	14742	ı	14802	ı	14862	ŧ	FIG.2-36	1
E L W R S Y G V E P A A V V G H S Q G E	ATCGCCGCCGCCACGTCGCCGGGGGCGCTCACGTTGGAGGACGCGGCGAAGCTCGTCGTG ++++++	IAAAHVAGALTLEDAAKLVV	GGCCGCAGCCGCCTGATGCGGTCGCTCTCCGGGGAGGGCGGCATGGCCGCCGTCGCGCTG	GRSRLMRSLSGEGGMAAVAL	GGCGAGGCCGCGGTGCGCGAGCGCCTGCGGCCGTGGCAGGACCGGCTCTCGGTGGCCGCG	G E A A V R E R L R P W Q D R L S V A A	GTCAACGGTCCCCGGTCGTGGTCTCCGGCGAGCCCGGCGCGCGC	VNGPRSVVVSGEPGALRAFS	GAGGACTGCGCGGCCGAGGGCATCCGCGTCCGCGACATCGACGTGGACTACGCCTCGCAC	E D C A A E G I R V R D I D V D Y A S H
•	14623		14683	-	14743	•	14803		14863	

	100 miles													
14982	1	15042	1	15102	ı	15162	<u> </u>	15222	ı	FIG.2-37	1			
TCGCCGCAGATCGAGCGGGTCCGCGAGGAACTCCTCGAAACGACCGGCGACATCGCGCCG	SPQIERVŘEELLETTGDIAP	CGCCCGGCGCGGGTGACGTTCCACTCTGTGGAGTCGCGGTCTATGGACGGCACCGAG	RPARVTFHSTVESRSMDGTE	CTGGATGCCCGGTACTGGTACCGCAACCTGCGCGAGACGGTGCGCTTCGCCGACGCCGTG	L D A R Y W Y R N L R E T V R F A D A V	ACGCGGCTGGCGGAGTCGGGATACGACGCGTTCATCGAGGTCAGCCCCGCATCCGGTCGTG	ж 	GTCCAGGCCGTCGAGGCGGTCGAAGAGGCTGACGGTGCCGAAGACGCGGTCGTAGTC	V Q A V E E A D G A E D A V V V	GGCTCGCTGCACCGCGACGGCGGTGACCTCTCGGCCTTCCTGCGGTCGATGGCCACCGCG	G S L H R D G G D L S A F L R S M A T A			
14923		14983		15043		15103		15163		15223) 			
	SUBSTITUTE SHEET													

										-38
15342	į	15402	ı	15462	1	15522	1	15582		FIG.2-38
CACGTGTCCGGTGTGGACATCAGGTGGGACGTCGCTCTGCCCGGCGCGCGC	HVSGVDIRWDVALPGAAPFA	CIGCCGACGIATCCGTICCAGCGCAAGCGCTACIGGCTCCAGCCCGCCGCACCCGCCGCCGCCGCCGCCGCCGCCGCC	LPTYPFQRKRYWLQPAAPAA	GCCTCCGACGAGCTGGCCTACCGCGTTTCCTGGACTCCGATCGAAAAGCCGGAGTCGGAA	ASDELAYRVSWTPIEKPESG	AACCTGGACGGCGACTGGTTGTCACACCCCTCATCAGTCCGGAGTGGACGGAAATG	N L D G D W L V V T P L I S P E W T E M	CIGIGCGAGGCCATCAACGCCAACGGTGGCAGGGCGTTGCGCTGCGAGGTGGACACGTCC	LCEAINANGGRALRCEVDTS	GCTTCGCGCACTGAGATGGCCCAGGCCGTCGCACAGGCCGGAACGGGATTCCGGGGCGTG 83+++++++-
15283		15343		15403		15463		15523		15583

					4	10/89				02.0	10.2 - 33 002
ı	15702	į	15762	ı	15822	1	15882	i	15942	ر ا	16
A S R T E M A Q A W A Q A G T G F R G V	CTCTCGTTGCTGTCGTCGGACGAATCCGCCTGCCGTCCGGGGGTTCCTGCCGGTGCGGTC	LSLLSSDESACRPGVPAGAV	GGCCTGCTCACCCTGGTCCAGGCGCTGGGCGATGCCGGGGTCGACGCACCGGTGTGGTGC	G L L T L V Q A L G D A G V D A P V W C	CTGACCCAGGGTGCGGTCCGCACTCCCGCCGACGACGACCTCGCCCGGCCTGCGCAGACC	LTQGAVRTPADDDLARPAQT	ACCGCGCACGGCTTCGCGCAGGTCGCCGGGCTGGAGCTGCCGGGCCGCTGGGGCGGTGTG	TAHGFAQVAGLELPGRWGGV	GTCGACCTGCCCGAATCGGTCGACGCGCGCGCTGCGTCTGCTCGTGGCAGTCCTGCGC	V D L P E S V D D A A L R L L V A V L R	GGCGGCGGCCGTGCCGAGGACCACCTCGCGGTCCGGGACGGCCGCCTCCACGGCCGTCGC
-	15643		15703		15763		15823		15883		15943

ATCAGCGAAGTCGCACCGGAGGAGTTCGCCGAGACGATCGCGGGCCAAGACCGCGTTGCTC

					41	1/89			-40	
ı	16062	1	16122	i	16182	ı	16242	, <u>.</u>	FIG.2-40	t
G G G R A E D H L A V R D G R L H G R R	GTCGTCCGCGCAAGCCTGCCGCAGTCCGGCTCGCGGAGCTGGACCCCCGCACGGGACCGTG	V V R A S L P Q S G S R S W T P H G T V	CTGGTCACCGGCGCGGCGAGCCCCGTCGGCGACCAACTGGTGCGGTGGCTCGCCGACCGG	LVTGAASPVGDQLVRWLADR	GGAGCCGAGCGGCTGGTGCTGGCCGGAGCCTGTCCGGGCGACGACCTGCTGGCCGCGGTC	G A E R L V L A G A C P G D D L L A A V	GAGGAAGCGGGCGCATCGGCCGTCGTGTGCGCCCCAGGACGCGGCGGCGCCTGCGCGAGGCG	E E A G A S A V V C A Q D A A A L R E A	CTCGGCGACGAGCCGGTGACCGCGCTCGTGCACGCCGGAACCCTGACGAACTTCGGCAGC	LGDEPVTALVHAGTLTNFGS
	16003		16063		16123		16183		16243	

					•	42/89			,	4	
16362	1	16422	1	16482	1	16542	ı	16602	į.	FIG.2-4	I
+++++++	VAPEEFAETIAAKTALL	GCCGTGCTGGACGAAGTCCTCGGCGACCGGGCCGTCGAGCGGGAGGTCTACTGCTCGTCG	DEVLGDRAVEREVYCSS	GTCGCCGGGATCTGGGGCGCCCCGGGATGGCCCGCCTACGCGGCAGGCA	IWGGAGMAAYAAGSAYL	GACGCGCTGGCCGAGCACCACCGCGCGGGGCCGCTCGTGCACCTCGGTCGCCTGGACG	AEHHRARGRSCTSVAWT	CCGTGGGCGCTGCCGGGCGGGCGGTGGACGACGGCTACCTGCGGGAACGCGGACTGCGC	LPGGAVDDGYLRERGLR	AGCCTCTCCGCCGACAGGGCGATGCGCACCTGGGAGCGGGTGCTGGCCGCCGCCGGGCGGTG	ADRAMRTWERVLAAGPV
 - 	S E	GCCGTGC	> 1	TCGCCG	છ	GACGCGC	A	CGTGGG	W	AGCCTCTCCGC	. 1 8
16303 -	H	G 16363 -	æ	G 16423 -	>	G 16483 -	Q	C 16543 -	Δı	A 16603 -	S

								-	4	
16722 -	16782	ı	16842	ı	16902	ı	16962	, (r IG.2-4	1
TCGTCGCGGTGCCGACGTGCCGGTGCTCAGCGAAGGCTTCGCCGCCACCCGG++++++	CCGACCGCGCTGTTCGCCGAACTCGCCGGCGGCGGGGGGGG	PTALFAELAGRGGQAEAEPD	AGCGGACCGACCGGCGAGCCGGCACACGGCTCGCGGGGCTTTCCCCCGGACGAGCAGCAG	S G P T G E P A Q R L A G L S P D E Q Q	GAAAACCTGCTCGAACTCGTCGCGAACGCGGTTGCCGAGGTGCTTGGCCACGAGTCCGCC	ENLLELVANAVAEVLGHESA	GCCGAGATCAACGTGCGCCGCGCGTTCAGCGAGCTCGGACTCGACTCGCTCAACGCGATG	A E I N V R R A F S E L G L D S L N A M	GCCCTGCGCAAGCGCCTGTCGGCGAGCACCGGCCTGCGGCTGCCCGCGTCGCTGGTTC	A L R K R L S A S T G L R L P A S L V F
16663	16723		16783		16843		16903		16963	
	·					S	UBSTI	TŲT	E SHE	ET

					4	4 / 8 9			7) t	-
17082	1	17142	ı	17202	ı	17262	ı	17322	- 510 2-43	17	ı
GACCACCCCACCGTCACCGCGCTCGCGCAGCACCTGCGCGCCCCGGCTCGTCGGTGACGCC	DHPTVTALAQHLRARLVGDA	GACCAGGCCGCGGTGCGCGTCGTCGGCGCCGACGAGTCCGAGCCCATCGCCATCGTC 17083+++++	D Q A A V R V V G A A D E S E P I A I V	GGCATCGGCTGCCGTTTCCCCGGCGGCATCGGCTCGCCCGAGCAGTTGTGGCGGGTGCTG	GIGCRFPGGIGSPEQLWRVL	GCCGAGGGCGCGAACCTCACCACCGGCTTCCCGGCCGACCGGGGCTGGGACATCGGGCGG	AEGANLTTGFPADRGWDIGR	CTCTACCACCCGGACCCGGACAACCCCGGCACCAGCTACGTGGACAAGGGCGGGTTCCTC	LYHPDPDNPGTSYVDKGGFL	ACCGACGCGGCGATTTCGACCCGGGCTTCTTCGGCATCACGCCCCGCGAAGCGCTGGCG	T D A A D F D P G F F G I T P R E A L A
			SU	BSTIT	UTE S	HEET					

					, 0	7 63			Ċ	FIG. Z-4 742
17442	1	17502	ı	17562	1	17622	1	17682	ī	F 16
ATGGACCCGCAGCACCCTCATGCTGGAGACGGCGTGGGAGGCAGTGGAACGCGGGGC	M D P Q Q R L M L E T A W E A V E R A G	ATCGACCCCGACGCCCTGCGAGGCACCGACACCGGCGTCTTCGTCGGCATGAACGGCCAG	I D P D A L R G T D T G V F V G M N G Q	TCCTACATGCAGCTGCTGGCCGGTGAGGCCGAACGCGTCGACGGCTACCAGGGCCTCGGA	SYMQLL'AGEAERVDGYQGLG	AACTCCGCGAGCGTGCTCTCCGGGCGCATCGCCTACACCTTCGGCTGGGAGGGCCCGGCG	NSASVLSGRIAYTFGWEGPA	CTGACGGTGGACACCGCGTGCTCGTCCTCGCTGGTCGGCATCCACCTCGCGATGCAGGCG	LTVDTACSSLVGIHLAMQA	CIGCGGCGCGGTGAGTGCTCCCTGGCGCTGGCCGGCGGCGTCACGGTCATGTCCGACCCG
17383		17443		17503		17563		17623		17683
		6		TITI ITE	CHE	ET				

					46/8	39		777	1 1	
t	17802	t	17862	ı	17922	1	17982	ו כו	H	1
G V T V M S D P	AGCGCGGGCTCGCCTCCGACGGTCGCTGCAAGGCG	A S D G R C K A	CGCGCTGTCGGAAGGCGTCGCCGCGCGTGGTGCTGGAG	GVAALVLE	TGCTGGCCGTGCTGCGCGGCAGCGCG	L A V L R G S A	CTCCCAACGGCCCGTCGCAGGAGCGG	PNGPSQER	CGGGCGTGCCGGCGGGACGTCGACGTCGTGGAG	AADVDVVE
LRRGECSLALAG	TACACCTTCGTCGACTTCAGCACGC	Y T F V D F S T Q R G L	TTCTCCGCGCGGCCGACGGCTT	FSARADGFALSE	CCGCTTTCCCGGGCGCGCGCCAACGGGCACCAGG	P L S R A R A N G H Q V	GTCAACCAGGACGGTGCCAGCAACGGTCTCGCCGCTCCCAACGGCCGGTCGCAGGAGCGG	V N Q D G A S N G L A A	GTGATCCGGCAGGCGCTCGCCGCTT	V I R Q A L A A S G V P
	17743		17803		17863		17923		17983	

	•				4	7/89			•	140	
18102		18162		18222		18282		18342	(L	FIG.2-45 18402	ı
	I EH		н		ا ن		ا ق	GCCGTGGAGGTGCTGCGCGAGGAGGTGCCGTGGCCGGCGGGTGAGCGCCCCCCGGCGGGGG	4 .	GGGGTGTCGTCCTTCGGCGTCAGCGGAACCAACGCGCACGTGATCGTCGAAGAGAGGCACCA ++++++	ים
GCGCACGGGACGGGCACCGAGCTCGGCGACCCGATCGAGGCCGGCGCGCTCATCGCGACC	A I	TACGGCCAGGACCGCGACCGGCCGCCTCGGCTCGGTGAAGACCAACATCGGCCAC	ა ა	acccaggccgcggcggcgcgcgggcgtgatcaaggtcgtgctggcgatgcggcacgg	H	ATGCTGCCCCGGTCGTTGCACGCCGACGAGCTGTCCCCGCACATCGACTGGGAGTCGGGG	S		æ	GCA	Æ
ATC	· H	ATC	Н		~	GAG	ចា	990	~	GAG	ធ
CIC	H	AAC	z	ATG	Σ	TGG	3	222	Дı	GAA	ы
-+	K	ACC.	E	GCGAT +	Æ	GAC +	Ω	200005	æ	GTC +	>
299	O	AAG	쏘	CIG	IJ	ATC	н	GAG	Ш	ATC	н
rcgaggccg	Æ	GTG	>	GTG	>	SCAC	I	CCGGCGGGT	ပ	GTG	>
GAG -+-	ចា	TCG-+-	ဟ	GTC	>	000	⊶.	3600	Æ	SCAC	H
ATC	н	0993	Ŋ	AAG	×	STCC	တ	3000	Ωı	2000	A
000	<u>α</u>	SCTO	H	SATC	н	SCTC	7	GTGG	Δ. 3	CAA(z
SCGACC	Ω	TGCGGCT	æ	CGT(>	GCCGACGAGCT	ធ	SAGGTGCCG	Сı	AAC	Ħ
	Ŋ	GCT	ы	 	ပ	CGA	Ω	GGT	>		U
GCT	н	225	Δι	 CGC	æ	292	Æ	GGA	ជា	CAG	ഗ
CGA +	ធ	5 + +	ĸ	+	Æ	GCA:+	Ħ	CGA +	យ	GGCGTCAGCGGAACCAACGCGCACGT	>
CAC	H	CGA	Ω	56660	Ŋ	GTT	H	525:	64	5500	Ŋ
999:	Ŋ	5001	ĸ	999	Ø	GGTC	S	rgcı	П	CTT	िंग
CGGGACG	H	GGA	Ω	2002	Ø	2007	æ	4GG7 	>	GTC	S
9901	ტ	CCAG	ø	999	¥	25.		GGA(ជា	GTC	ഗ
SGCA	Œ	VCG	ტ	ACCCAGGCCGC	ø	ATGCTGCCCCGGTCGTTC	н	2061	>	GGGGTGTCGTCCT	>
	æ		×		Ħ		Σ		Æ		G.
18043		18103		18163		18223		18283		18343	

	18403		AGA	gagcagga +	GGA	299	+:	+	CAC	CGA	AGCGC	553	J. J.	3CT(GCTGCC	STT	CGT	CGTGCTGT	37CC	GCAGAGCAGGAGGCCGCCGCACCGAGCGCGGTCCGCTGCCGTTCGTGCTGTCCGGCCGC +++++++		18462	
•		Ø	ជ	ø	দ্র	K	æ	K	H	ធ	¤	O	ď	H	Δı	ក្រ	>	ı	S	R	•		
	18463		CGA	AGC +	GCCGT -+	.661	2000:	+	CCA	299	500) 	GCT	CTCGCCG	CGA	SCA	+	3000	GAC	AGCGAAGCCGTGGTCGCGGCCCAGGCCCGCGCGCTCGCCGAGCACCTGCGCGACACCCCG		18522	
_		ഗ	ធ	4	>	>	A	Æ	ø	Ø	æ	æ	H	K	ជ	Ħ	ы	~	Ω	E E	•	1	
	18523		I CCI	20000	1001	GAC	, , , ,	+-	299	GTG	GAC -+-	GCT) (300	GAC	990	CAG	990	+- CGCG(3TT(GAGCTCGGCCTGACCGACGCGTGGACGCTCGCGACCGGCAGGGCGCGGGTTCGACGTG		18582	
		ធ	н	Ŋ	ц	H	D	Æ	Æ	×	H	н	Æ	Ħ	Ŋ	~	æ	æ	Įτί	ρ		ı	48
CHEE	18583		CGAGC	cccccrc	 		, , , ,	+-	CGA	500,	929292	999	CGT	GTG +	292	GGA	GCT	GGA(+	200	CGAGCCGCCGTGCTCGGCGACGACCGCGCGGGCGTGTGCGCGGGAGCTGGACGCGCGCTGGCC ++++++		18642	8/89
т		· ¤	Æ	Ø	>	1	ტ	Ω	Ω	ĸ	K	r	>	ပ	K	ഥ	H	Ω	Ø	I.		ı	
	18643			-+	200	GTC	3660	CG2) CGC	CGT	CGTCGCGC	225	CGGT	GAC	TGACCTCC	292	225	92922	CAA	GAGGGCCGCCCGTCGGCCGACGCCGTCGCGCCGGTGACCTCCGCGCCGCGCGCAAGCCGGTC		18702	
		回	ဗ	ĸ	Α.	S	K	Ω	Æ	>	æ	, or	>	H	S	Æ	Д	ĸ	×	ь ч		1	
	18703		CIGGICIICCCCGG	CTJ	χχ)555	 ?::::::::::::::::::::::::::::::::::	4GGC	3000	200	GTG-+	GCGCAGTGGGTCGGCATGGCAC	090	CAT	099	ACG	CGA	CGATCTG	GCT	CTGGTCTTCCCCGGCCAGGGCGCGCAGTGGGTCGGCATGGCACGCGATCTGCTGGAATCC	22	FIG.2-47 18762	4
		H	>	ĹĿ	Дı	. ტ	Ø	G	æ	ø	3	>	Ö	Σ	Ø	<u> </u>	Ω	ы	ы	<u>പ</u> ഗ		i	

					49/	89				FIG.2-48
18822	ı	18882	ı	18942	ì	19002	ı	19062	i	FIG.:
TCCGAGGTGTTCGCCGAGTCGATGAGCCGGTGCGCCGAGGCGCTCTCGCCGCACACCGAC	SEVFAESMSRCAEALSPHTD	TGGAAGTTGCTCGACGTCGTCCGCGCGACGGCGGTCCCGACCCGCACGAGCGCGTCGAC	W K L L D V V R G D G G P D P H E R V D	GIGCICCAGCCGGIGCICTICICGAICAIGGICICGCIGGCCGAGCIGIGGCGCGCGCAC	V L Q P V L F S I M V S L A E L W R A H	GGCGTGACCCCGGCCGCCGTCGTCGCCACTCGCAGGGCGAGATCGCCGCGCGCG	G V T P A A V V G H S Q G E I A A A H V	GCGGGCGCGCTGTCGCTGGAAGCCGCCGCGAAGGTGGTGGCCCTGCGCAGCCAGGTGTTG 19003+++++++	A G A L S L E A A A K V V A L R S Q V L	CGCGAGCTCGACGACCAGGGCGGCATGGTGTCGGTCGGCGCGCCGCCGCGCGCG

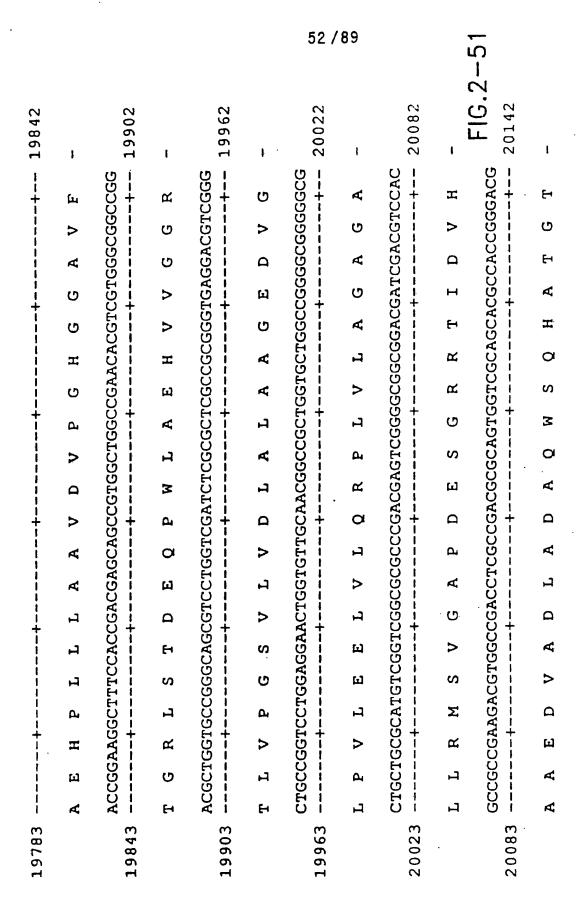
					50/	89				· ·	r .c.z – 49 482
ı	19182	1	19242	ı	19302	ı	19362	ı	19422	ا ا	۲ ۱ ۵.
RELDDOGGMVSVGASRDELE	ACCGIGCICCGCGCGCTGGGACGGCCGTGICGCGGTGGCCGCCGTGAACGGGCCTGGCACC	T V L A R W D G R V A V A A V N G P G T	AGCGTCGTTGCCGGGCCGACCGGGAGCTGGACGAGTTCTTCGCCGAGGCCGAGGCGGGGCGGG	SVVAGPTAELD'EFFAEARR	GAGATGAAGCCGCGCCGGATCGCCGTGCGCTACGCCTCCCACTCCCCGGAGGTGGCGCGC	EMKPRRIAVRYASPEVAR	ATCGAGGACCGGCTCGCGGCCGAGCTGGGCACCATCACCGCCGTGCGGGGCTCGGTGCGG	I E D R L A A E L G T I T A V R G S V P	CTGCACTCCACGGTGACCGGCGAGGTCATCGACACCTCCGCGATGGACGCCTCCTACTGG	LHSTVTGEVIDTSAMDAS·YW	TACCGCAACCTGCGCCGACCAGTGCTCTTCGAGCAGGCGGTGCGCGGTCTGGTCGAGCAG
	19123		19183		19243	- 11-	19303		19363		19423
			S	UBS	TITUTE	SHE	EI				

÷

÷

						51 /8	39		-50		
ı	19542	ı	19602	1	19662	ı	19722	ı	FIG.2-50	i	
YRNLRRPVLFEQAVRGLVEQ	GGCTTCGACACCTTCGTCGAGGTGAGCCCGCACCCGGTGCTGCTGATGGCGGTCGAGGAG	G F D T F V E V S P H P V L L M·A V E E	ACCGCCGAGCACGCGGGCGCGGAAGTCACCTGCGTGCCGACGCTGCGCCGCGCGCG	TAEHAGAEVTCVPTLRREQS	GGACCGCACGAGTTCCTGCGCAACCTGCTGCGGGCTCACGTGCACGGCGTCGGCGCGCGACGAC	GPHEFLRNLLRAHVHGVGAD	CTGCGTCCGGCGGTGGCCGGGGGACGGCCGGCCGAGCTGCCCACCTACCCGTTCGAACAC	LRPAVAGGRPAELPTYPFEH	CAGCGCTTCTGGCCGCCGCCGCCCGCCCGCCGACGTCTCGGCGCTGGGCGTGCGCGGG	QRFWPRPADVSALGVRG	GCGGAGCACCCGCTGCTCGCCGCGGTCGACGTGCCGGGCCACGGCGGTGCGGTTTC
	19483		19543		19603		19663		19723	•	

દ



2

GGGGCGGACGCGATGGCCCTGCGGGTCACCGACCCGGCAGGCCACCTGGTCGCCACGGTC ++	A G H L V A T V -	CAGCACCGGGGAGAAGTGGGAGCAGCCCGAACCGCGCGGTGGC +	- 99 x d a d O a	TCTGGACTGGGGACGGCTAGCCGAGCCCGGCTCGACCGGTCGT +	AEPGSTGR -	CTCGGACCTCGACGCCGTCCTGCGGTCCGGTGAACCCGAACCC	LRSGEPEP -	GACGCGGTCCTGGTCCGCTACGAACCCCGAAGGCGACGACCCCCCCC	DPRAAARH -	CGCGCTCGTGCGCCGCTGGCTCGAACAGGAGGAGCTGCCGGGC +
	i		1		I .		1		1	
GGT -+-	>	TGG -+-	ប	TCG +-	ĸ	ACC:+-		-+	Ħ	7666
CAC	Ħ	990	ტ	990	ტ	ZCGP	ជា	3000	ĸ	5555
292	Ø	909	æ	GAC	H	ACC	വ	3660	K	\GC1
GGT +		ACC +	ы	CTO:+	တ	TGA	闰	+ +-	K	+-
CCT	н	CGA	ធ	990	ບ	9903	ഗ	3000	Ø	
CCA		225.	Д	900	Д	GTC	တ	0000	α,	AAC?
AGG		GCA		4 500	ធ	929		VCCO		rcg.
999		GGA	ចា	AGC		1 1 1 1 1	J	CGA		;GC1
222	Д	GTG	3	GCT	н) 	>	3CGP	Ω	CTC
CGA	Ω	GAA	×	ACG	ĸ	0901	æ	AAGG	Ŋ	0005
CAC	H	GGA	ធ	TGGGGAC	ტ	1500 1+1	Ω	-+-	回。	+
.GGT	>	992	Ŋ	CTG	×	[]	h	AACO	Ф	rcg1
000	K	CAC	H	799;	Ω	7993	Ω	VCG7	ជា	, GC1
+ - +	н	CAG	S	TCI +	н	CTC +	လ	3CT?	×	+-)CCC(
3660	æ	GACTCGCTGGTCGTCCG	K	GAGGGCGAGCTGCACGC	æ	GTGGTCGCGGCCGATGC	K	2001	R.	GGCGTCCTCGGGCCGC
GAT	Σ		>	13GC	Ħ	7900	Ω	rGG1	>	3660
090	Ø	.GGT	>	1001 1111	H	0993	K	GGTCCT +	H	CTCTC
+	Ω	GACTCGCTGGT	H	CGAG	ជា	+	K	3663	>	Σς <u>Τ</u>
3660	K	CTC	ဟ	\GGG	Ŋ	166J	>	ACGC	K	3CG1
	ტ		Ω		ធ		>		Ω	
20503		20563		20623		20683		20743		20803
			CI	IDOTII	nite (SHEET	•			

					. 5	5/89)				-54
1	20922	ł	20982	ı	21042	1	21102	t	21162	1	FIG.2-54
; V L W A A A L V R R W L E Q E E L P G	GCGACGCTGGTCATCGCCACGTCCGGCGGTCACCGTGTCCGACGACGACAGCGTTCCC	TIVIATSGAVTVSDDDSVP	GAACCCGGCGCCGCGGATGTGGGGGCGTGATCCGCTGTGCGCAGGCCGAGTCGCGCGGAC	PGAAAMWGVIRCAQAESPD	CGGTTCGTGCTCCTCGACACCGACGCGGAACCTGGGATGCTGCCTGC	R F V L L D'T D A E P G M L P A V P D N	CCGCAGCTCGCGTTGCGCGGCGACGTCTTCGTGCCGCGCCTCTCGCCGCTCGCACCT	P Q L A L R G D D V F V P R L S P L A P	TCCGCGCTGACGCTTCCGGCAGGCACCCCAACGTCTCGTGCCGGGGTGACGGGGGGGG	SALTIPAGTORLVPGDGAID	TCCGTGGCCTTCGAGCCCGCACCCGACGTCGAGCAGCCGCTCCGGGCGGG
Ŋ	20863 -	æ	20923	ជ	20983 -	-	21043	~	21103		21163

						56/	89				55
	C)		~		7		2		2		FIG.2-55
	21282		21342		21402		21462		21522		FIG.: 21582
1	21	í	21	I		1		1		l.	
ሺ	CGGAGTCAACTTCCGCGACGTCCTCCTCGCACTCGGCATGTAT	×	CCGCAGAAGGCGGACATGGGCACCGAGGCCGCCGGTGTCGTCACGGCGGTCGGACCGGAC ++++++	Q	GTGGACGCCTTCGCGCCGGGAGACCGGGTGCTCGGCCTGTTCCAGGGAGCCTTCGCGCCGCG	Δı	ATCGCGGTCACCGATCACCGGCTCCTCGCACGAGTGCCGGACGGCTGGAGCGACGCCGAC	Ω	GCCGCGCCGTGCCCATCGCCTACACCACGGCGCGCATTACGCGCTGCACGATCTCGCGGGG	U	CTGCGCGCGGGTCAGTCGGTGCTCATCCACGCAGCGGCAGGCGGTGTCGGCATGGCGGCCC
>	SCA	Σ	3AC	ы	ICG	æ	ACG	Ø	7CG	A	11 GG
ធា	555	Ŋ	90	Ŋ	CT	Ĺμ	500	Q	ATC	귀	3CA
Ŋ	CACT	1	GG1 +	>	GGAGC +	æ	GGAG -+	လ	+ +	Ω) +
æ	292	4	360	Æ	999	r	CIG	3	GCA	I	TGT
ጁ	CCT	ы	CAC	H	CCA	α	 	ပ	GCT	ы	550
н	CGICCI	H	GCCGGTGTCGTCACGGCGGTCGG	>	GTT	[T4	GGA	; D	VCGC	K	AGG
գ	CGT 	>	TGT +	>	CCT	ы	÷	ሷ	TTT	×	999
a	CGA(Q	555	Ŋ	990	ဗ	AGT	>	GCA	H	AGC
ធ	Ď	~	292	Ø	GCT	ы	ACG	×	099	A	292
>	CTT -+-	ĮΞι	AGGCC	Æ	GGT -+-	>	292	Ø	CAC -+-	H	CCA -+-
Ω	CAA	z	CGA	回)))	æ	CCT	ы	CAC	Ħ	CAT
Ф	AGT	>	CAC	H	AGA 	Ω	GCT	H	CTAC	≯	GCT
æ	+ 990	Ŋ	+-	O	+	ບ	+ +	ĸ	CATCGCC	Ø	.GGT
Д	CAC	H	CAT	Σ	229	Δı	TCA	Ξ	CAT	H	GTC
ធា	252	K	GGA	Ω	090	A	CGA	Q	000	а	TCP
ĮΣι	909	æ	299	Ø	CTT	ជ្រ	CAC	H	CGT	>	999:
Æ	CGT +	>	GAA 	×	CGCCT	Æ	GGT 	>	999	K) 1
>	GTGGACGTGCGCGCCAC	Ω	CCGCAGAAGG	ø	GGA 	Ω	292	K	900000000000000000000000000000000000000	Ø	900 1 - 1
လ	GH	>	Ö	Д	GI	>	AT	н	Ü	Ø	S I
	21223		21283		21343		21403		21463		21523

					57 /	89			-56	•
	42		,02		21762		322		FIG.2-56	
i	21642	1	21702	i	21.	ł	2182	ı		t
SVLIHAAAGGVGMAA	GTCGCGCTGGCCCCGCCGAGCGGGGGGGGGTGTTGGCCACCGCCGGCCCGGCCCAAGCAC 	RAGAEVLATAGPAKH	GGGACGCTGCGGCGCTCGGTCTCGACGACGAGCACATCGCTTCCTCCCGGGAGACCGGT ++	LGLDDEHIASSRETG	TTCGCCCGGAAGTTCCGGGAGCGCACCGGAGGCCGCGGGGGCGTGGTGGTGCTCAACTCG	RERTGGRGVDVVLNS	CTCACCGGGGAACTGCTCGACGAGTCCGCGGATCTGCTCGCCGAGGACGGCGTCTTCGTC	LDESADLLAEDGVFV	GAGATGGGCAAGACCGACCTGCGGGACGCCGGGGACTTCCGGGGCCGATACGCCCCGTTC	D L R D A G D F R G R Y A P F
თ უ	3000	K	3660	Æ	AGT1	ĮΞı	AAC	ı	AGA(Ħ
Ŋ	000000000000000000000000000000000000000	Æ	IGC(+	L R	GGA.	×	CACCGGGGAAC	ជ	GGGCA	×
A	200	П) C	A	000	E G	TGG	G
L R	GTCG(V	GGGACGCTGCGG +	E G	TTCGCCCGGAAG	F	CTCA	L T	GAGAT	Σ ω΄
1	21583		21643	J	21703	144	21763		21823	_

					5	8/89		•		ည်
21942	ı	22002	1	22062	ı	22122	ı	22182	į	FIG.2-5
GACCTCGGCGAGGCGGGTGACGACCGGCTCGGGGAGATCCTGCGCGAGGTCGTCGCCTG	D L G E A G D D R L G E I L R E V V G L	CTGGGCGCCGGGGAGCTCGACCGGCTCCCGGTATCGGCGTGGGAGCTGGGATCCGCGCCCC	LGAGELDRLPVSAWELGSAP	GCGGCGTTGCAGCACATGAGCCGGGGCAGGCACGTCGGCAAGCTCGTGCTGACCCCAGCCC	A A L Q H M S R G R H V G K L V L T Q P	GCGCCGGTGGACCCGGACGGCACGGTGCTGATCACGGGTGGCACCGGCACGGTTCGGACGG	APVDPDGTVLITGGTLGR	CIGCICGCGCGCCACCICGICACCGAGCACGGCGTGCGGCACCIGCIGCTGCTGGTCAGCAGG	LLARHLVTEHGVRHLLLVSR	CGCGGCGCGCGCGCCCGGGTTCCGACGAGCTGCGCGCGGAGATCGAGGACTTGGGCCGCG
21883		21943		22003		22063		22123		22183

						59	/89				-58
ı	22302	ı	22362	1	22422	ı	22482	ı	22542	1	FIG.2-
RGADAPGSDELRAEIEDLGA -	TCCGCGGAGATCGCGGCTTGCGACACCGCCGACGCGCGCTTTCGGCGCTGCTGGAC	SAEIAACDTADALSALLD	GGGCTGCCCCGGCCGCTGACCGGTGTCGTGCACGCGGCGGGTGTGCTGCCTGGCCGACGGGCTG	G L P R P L T G V V H A A G V L A D G L	GTCACCTCCATCGACGAGCCGGCGGTGGAGCAGGTGCTGCGCGCCCAAGGTCGACGCGGCG	V T S I D E P A V E Q V L R A K V D A A	TGGAACCTGCACGAGCTGACCGCGAACACCGGTCTGAGCTTCTTCGTGCTGTTCTCGTCCTGTTCTCGTCC	WNLHELTANTGLSFFVLFSS	GCGGCGTCGGTGCTAGCCGGCCCCGGGGCAGGGCGTGTACGCGGCCGCGAACGAGTCGCTC ++++++	A A S V L A G P G Q G V Y A A A N E S L	AACGCGCTGGCTGCCCTCCGGAGGACGCGCGGCCTTCCCGCGAAGGCGCTCGGATGGGGA ++++++
	22243		22303	-	22363	-	22423		22483		22543

						60/8	19			L	FIG. 2-59
										(
	22662		22722		22782		22842		22902		FIC 22962
l		1		1		1		1		ı	
ტ	CTGTGGGCGCAGGCCAGCGAGATGACCAGCGGACTCGGCGACCGCATCGCCCGGGGGGGG	ტ	GTCGCCGCGCTGCCGACCGAGCGGGCGCTCGCACTGTTCGACAGCGCCCTGCGCCGCGGC	Ŋ	GGTGAGGTCGTGTTCCCGCTGTCCATCAACCGTTCCGCGCTGCGCAGGGCCGAGTTCGTG	>	TGGTCAGGGCGAAGCTGCGCGCCGCCGGGCAGGCCGAGGCG	Æ	ACCGGCTCGCTCGGTCCGAGTCCGACCAGGTCGCCGGG	Ŋ	CACACGCGGCGGTCTCCGGGTACGGCTCGGCCGACCAG
3	GAC	H	900	æ	GTT	ĹĿı	CGA	ы	292	æ	CGA
Ŋ	500	æ	929,	ĸ	CGA	ជ	0991	æ	\GGT	>	3660
ы	ATCGC	Ø	+-+	П	CAGGGC	æ	GCCGCCGGGCAGG	Ø	ACC.	a	CGGCTCG
Æ	CAT	H	292	Ø	CAG	ĸ	990	G	,CG2	Д	ACGO
×	900	ĸ	CAG	S	929,	ĸ	090	æ	GTC	လ	GTP:
æ	CGA	Ω	CGA	Ω	GCT	H	3000	Æ)CG2	ធ	CTCCGC
Δι	2992	Ŋ	GTT +	Įτι	TCCGCG	æ	AGGGCGAAGCTGCGC	ĸ	3GTC	ഗ	ICTC
ы	ACT	ы	ACT	7	STTC	လ	AGCJ	IJ	STCC	ĸ	
Ŋ	9906	Ŋ	, , ,	K	CAACCG-+	K	CGAJ	×	5000	ტ	
r.	CCAG +-	တ	CGCTCG	H	rcaz	Z	3660	A	FCG(Æ	+ 52555
Ħ	GAC	H	3666	A	CCAT	Н	rcac	K	3903	h	ACG(
ĸ	AGAT	Σ	4GC0	R	TGT(S	TGG7	>	ACC(æ	CAC
x	CAGCGA	ជា	;cg;	臼	+ +	Н	3CA:	Σ	rcG;	Ω	GTT(-+-
1	CAC	လ	GAC	H	TCCCGC	Д	GCGGCA	ტ		>	rgc(
Æ	099	A	000	Д	GIJ	ĮΉ	000	ĸ	, CG	>	
Æ	GCA	ø	CCGCGCTGCCGACCG	7	CGT	>	CC	1	GAA.	Z	AACT
ы	999	æ	000	A	AGGTCG	>	GGTC-+	>	3600	ы	1907
Æ	GTG	3	GTCGC	Ø	TGA	Ш	CCGGAGGTCCTGCGCGGCA	កោ	GCAGGGCCGAACGTGGTCG	Q	CTGGCCGAACTGGTGCGTT
Z	C	н		>		Ŋ		а		Æ	
	22603		22663		22723		22783		22843		22903

CÉGGCGGACGCCCCGAGCACGTCCGCGATCAGCGAGGACGCCAGTGACGACGAGCTGTTC

					61/	89		(9-	
								(FIG.2-60	
	23022		23082		23142		23202		F10	
t	23(1	23	t		1		ì		1
Ö Q	CTCCCCGAGCGCAAGGCGTTCAAGGACCTCGGTTTCGACTCGCTGGCCGCGGGTGGAGCTG ++++++	고 일	GCGACCGGCGTGCGCTGCCCAGCACGTTGGTGTTCGACCAC	Д	GCCGAACACCTGCGGGACAGGCTGTTCGCGGCCTCACCGGCG	Ч	GTGGACATCGGCGACCGGCTGGACGAGCTGGAGAAGGCGCTCGAAGCCCTTGTCCGCCGAG ++++++	ы «	GACGGGCACGACGTGGGCCAGCGCCTGGAGTCGCTGCTGCGCCGGTGGAACAGCAGG	S R
I A	TEG	>	FTC	Eu.	ICA	S	TCC	S	AAC	z
	990:		3TG.	>	2005	Æ	DEC.	ы	7GG	3
S S	AAGGACCTCGGTTTCGACTCGCTGGCCGCGG	A A	TTGC-+	ı	.+ :ece(Æ	+	Æ	+	در
ъ	SCTG	ı	CACG	H	OFF	[z.	CGAP	ធ	929	æ
U) I C	S	CAG(ഗ	3CT(H	GCT	Ţ	GCT	ы
ග	GAC	Q)) + -	۵.	CAG(æ	÷	A	GCT 	IJ
>	TTC	[Li	SCT(ы	3GA(Ω	GAA	×	GTC 	က
Æ		ტ	3000	æ	300	æ	GGA	ជា	GGA	ា
K)CTC	H	3GT(>	CCT(H	GCT -+-	7	CCT -+-	Ы
æ	SGAC	Ω	2660	ပ	ACA(Ħ	CGA	ធ	 929	ĸ
Ħ	CAAC	×	GACC	۲	CGA	ப	GGA(Ω	CCA	a
တ <u>-</u>	GTT(Ĺų	CGC +	A	GGC +	Æ	GCT +	H	+ 999	ტ
K	3GCG	æ	CGCAACCGCCTCGGTACC ++	Ħ	CCGACTCCGCTGGCGGTG ++	>	5522	ぬ	CGT	>
>	CAA(×	595	Ŋ	299	Ø	CGA	Ω	CGA	Ω
ч	30G(~	CCT	H	GCT	1	CATCGG	ပ	CGA	Ω
មា	CTCCCCGAGCGCAAG	ជ	CGCAACCGCCTCGGT	ц	CCGACTCCGCTGGC	Д	CAT +	Н	GACGGGCACGA	Ħ
K	Ö	а	CAA	z	GAC	E→	GGA 	Ω	555	Ŋ
н	CTC	H		æ	_	凸		>		Ω
	22963		23023		23083		23143		23203	

					6	2/89			(FIG. 2—67 622	
23322	ī	23382	1	23442	ı	23502	ı	23562	C L	F16.2	I
+	ADAPSTSAISEDASDDELF	TCGATGCTCGACCAGCGGTTCGGCGGGGGGAGGACCTGTAGATGAGCGGTGACAACGGC	M L D Q R F G G G E D L * M S G D N G	ATGACCGAGGAAAAGCTCCGGCGCTACCTCAAGCGCACCGTCACCGAGCTCGACTCGGTG ++++++	TEEKLRYLKRTVTELDSV	ACCGCGCGCCTGCGTGAAGTCGAGCACCGGGCCGGTGAGCCGATCGCGATCGTCGGCATG	ARLREVEHRAGEPIAIVGM	GCGTGCCGGTTCCCCCGGCGACGTGGACTCGCCGGAGTCGTTCTGGGAGTTCGTGTCCGGCC	CRFPGDVDSPESFWEFVSG	GGCGGGGACGCCATCGCGGAGGCCCCCGCCGACCGCGGCTGGGAGCCGGACCCCGACGCG	G D A I A E A P A D R G W E P D P D A
23263 -	α .	T 23323 -	S	A 23383 -	Σ	A 23443 -	. H	23503 -	¥	23563 -	v

23682		23742		23802	63,	789		23922	(FIG.2-62	
TGCTCGCGGCCGCGGCGACTTCGACGCGGGCTTCTTCGGGATCTCG	MLAAAGDFDAGFFGIS -	CCGCGCGAGGCGCTGGCGATGGACCCGCAGCGGGATCATGCTGGAGATCTCGTGGGAG ++	LAMDPQQRIMLEISWE -	GCGCTGGAGCGCGCCGGCCACGATCCGGTGTCCCTGCGCGCGGCAGCGCGACCGGGGTGTTC ++++++	AGHDPVSLRGSATGVF -	TACGGCCCGCGACCGACGAGGCCCCGGACGAGGTCCTG	T V D Y G P R P D E A P D E V L	CGCCTCCAGCGTCGCCTCCGGCCGGGTCGCCTACTGCCTG	TGTASSVASGRVAYCL -	rgaccgtcgacaccgcctgttcctccgggctcaccgccctg	PAMTVDTACSSGLTAL -
CGGCTGGGCGGGA	R L G G M	CCGCGCGAGGCGC	P R E A]	GCGCTGGAGCGC(A L E R	ACCGGTGTCGGCACCGTGGAC	T G V G	GGCTACGTCGGCACCGGCAC	. 5 A X 5	GGCCTGGAAGGCCCGGCGA	G I E G
23623		23683		23743		23803		23863		23923	

		0476	,		9
24042	24102	24162	24222	0 1	FIG.2-6
CACCTGGCGATGGAGTCGCTGCGCCGGGACGAGTGCGGCCTGGCGCTGGCCGGCGGCGTG 23983+++++++	ACGGTGATGAGCAGTCCCGGGGCGTTCACCGAGCCAGGCGGGCG	GACGGCCGCTGCAAGCCGTTCTCGAAGGCCGCCGACGGGTTCGGCCTGGCCGAGGGTGCC 24103+++++++	GGGGTCCTGGTGCTGCAACGGCTGTCGGCCGCGGCGGGGGGGCAGACCGGTGCTGGCC 24163+++++++	GTGCTGCGGGGCTCGGCGCTCAACCAGGACGGCCCAGCAACGGGCTGACCGCGCCGAGC 24223+++++++-	GGACCCGCGCAGCAGCGGGTCATCCGCCGGGCGCTGGAGAACGCCGGTGTCCGGGCGGG
				ROTITUT	FSHE

					6	5/89					64
1	24402	1	24462	ı	24522	ı	24582	l	24642	1	FIG.2-64
PAQQRVIRALENAGVRAG -	GACGTCGACTACGTGGAGGCCCACGGCACCGGCACCAGGCTGGGCGACCCCATCGAGGTG	V D Y V E A H G T G T R L G D P I E V	CGCGCTGCTCTCGACCTACGGCGCGGAACGCGACCCGGACGATCCACTGTGGATCGGT	A L L S T Y G A E R D P D D P L W I G	TCGGTCAAGTCCAACATTGGCCACACCCCAGGCCGCCGCCGGCGTCGCCGGGGGTGAAG	V K S N I G H T Q A A G V A G V M K	GCGGTGCTGGCGCTGCGGCACGCGCGAGATGCCGCGCGCACGTGCACTTCGACGAGCCCTCG	V L A L R H G E M P R T L H F D E P S	CCGCAGATCGAGTGGGACCTGGGCGCGGTGTCGGTGGTGTCGCAGGCGCGGGGGGTGGTCCC ++++++	QIEWDLGAVSVVSQARSWP	GCCGGCGAGAGGCCCCCGCAGGGCGGCGTCTCCTCGTTCGGCATCAGCGGCACCAACGCG
<u>ග</u>	G. 24343 -	Ω	CA 24403	Ħ	T 24463 -	တ	. G 24523 -	Æ.	C 24583 -	<u>،</u>	. G 24643 –

					66	6/89				1	-65
1	24762	•	24822	1	24882	1	24942	į	25002	1	FIG.2-65 25062
GERPRAGVSSFGISGTNA	CACGTCATCGTCGAAGAGGCGCCCGAGGCCGACGAGCCCGAGCCGGCACCCGACTCGGGT ++++++	IVIVEEAPEADEPEPAPDSG	CCGGTCCCGCTGGTGTTGTCCGGCCGCGACGAGCAGGCGATGCGGGCGCGCAGGCGGGACGG	VPLVLSGRDEQAMRAQAGR	CTGGCAGACCACCTCGCCCGCGAGCCGGGAACTCGTTGCGCGACACCGGTtTCACGCTG +++++++-	LADHLAREPRNSLRDTGFTL	GCCACCCGCCGCAGCGCGTGGAGCACCGCGCGGTGGTGGTCGGCGACGCGCGACGACGCCC	A T R R S A W E H R A V V G D R D D A	CTCGCCGGGCTGCGCGCGGTGGCCGACGGCCGCATCGCCGACCGGACGGCCACCGGGCAG	LAGLRAVADGRIADRTATGO	GCCCGAACTCGCCGCGGCGTCGCGATGGTGTTCCCCCGGCCAGGGCGCGCAGTGGCAGGGG
4	24703 -	Ħ	24763 -	д	24823 -		24883	7	24943]	25003

					67/89)			FIG.2-66	
							6.1		7	
	25122		25182		25242		25302		F1G.	
1		ı		1		1		I	14	ı
ပ	GCGGGAGTCGCAGGTATTCGCCGACTCGATCCGCGACTGCGAG ++	ធ	CGTCGACTGGTCGCTGACCGACCTGCTCAGCGGCGCGCGC	വ	CGTCCAGCCCGCGCTCTTCGCCGTCATGGTGTCGCTGGCGGCG	A	CGTCGAGCCCGCCGCGGTCGTCGGCCACTCGCAGGGCGAGATC ++++++++	н	GCCGCCGCGCACGTCGCCGCGCGCTCACCCTGGAGGACGCCGCCAAGCTCGTCGCGGTC	>
ø	CTG	ပ	9 9 1	æ	255	æ	CGA	Ш	000	Æ
3	CGA(Ω	000	æ	GCT	ı	999	U	CGT	>
ø	GATCCGCG	ĸ) 1 1 1 1 1	Ŋ	3TC(လ	3CA +	ø	GCT +	ħ
æ	SATC	н	CAGO	တ	3GT(>	DHO I	ဟ	CAA	쏘
ပ	TCC	S	SCTC	H	CAT	Σ	CCA	Ħ	292	æ
a	CGCCGACT	Ω	CCTC	H	CGCCGT(>	CGG(Ŋ	CGGCGCGCTCACCCTGGAGGACGCCGCCAAGCTCGT	Æ
O) +	Æ	CGA 	Ω	+ CCC	æ	CGT +	>	.GGA 	Ω
Д	ATT	[z _i	GAC	E	CTT	Ē	GGT	>	GGA 	ជ
Ĭщ	3GT/	>	SCT	н	GCT	ı	292	Æ	CCT	H
>	rcgcaggt +	O	67.C	လ	GTCGTCCAGCCCGC	A	-+- 292	Æ	CAC -+-	H
Σ	OH I	S	CTG	3	000	۵	225	<u>م</u>	GCT	ы
Æ	3GA(ធ	CGA	Ω	CCA	ø	CGA	ធ	 292	æ
>	3CG +	ĸ	CGT	>	CGT	>	CGT +	>	+ +	Ŋ
Ŋ	SCT	н	GCA	Ħ	CGT	>	 	Ö	252	A
æ	CCT	1	001	Д	CGAC	Ω	CCA	I	CGT	>
ď	CGA(Q	000 	æ	GGT	>	CTC	တ	GCA	I
H	+	ĸ	# - +	1	Ü +	ದ	+ 909 +	æ	+ CCC	æ
ĸ	ATGGCCCGCGACCTGCT	Ø	CGGGCGCTGGCCCCGCA	æ	CTGGACCGGGTCGACGT	Ω	CTGTGGCGCTCCCACGG	Z	GCCGCCGCGCAC	4 .
Æ	AT 	Σ	CG	ፈ	CH	ı	E	1	00 1	Æ.
	25063		25123		25183		25243		25303	

					68/8	39			(F 1 6.2-6 7
25422	ı	25482	ı	25542	Į.	25602	ı	25662	ı	۲۱ ن. 25722
3GCTCGGCGGCCAGGGCGGCATGGCGTCGTTCGGGCTG	RSRVLARLGGQGGMASFGLG	ACCGAGCAGGCGGCCGAACGGATCGGGCGCTTCGCGGGCGCGCGC	TEQAAERIGRFAGALSIASV	AACGGCCCCCGGTCGTCGTCGTCGCGGGGAGAGCGGGCCGCTGGACGAGCTGATCGCC	NGPRSVVVAAGESGPLDELIA	GAGTGCGAGGCCGAAGGCATAACGGCGCGCCGCATCCCCGTCGACTACGCCTCCCACTCA 25543+++++++-	ECEAEGITARRIPVDYASHS	CCGCAGGTGGAGTCGCTGCGCGAGGAGCTGCTGACCGAGCTGGCGGGCATCTCCCCGGTG	PQVESLREELLTELAGISPV	TCGGCGGACGTGGCGCTCTACTCGACCACGACCGGGCAGCCCATCGACACCGCCACGATG
25		25		8		2.			MP1 17	7 7

					69	9/89					FIG.2-68
1	25782	ı	25842	i	25902	ı	25962	1	26022	1	26(
ADVALYSTTTGOPIDTATM	GACACCGCCTACTGGTACGCGAACCTGCGCGAGCAGGTCCGCTTCCAGGACGCGACGCGGG	TAYWYANLREQVRFQDATR	CAGCTCGCCGAGGCGGGTTCGACGCGTTCGTCGAGGTCAGCCCGCATCCGGTGCTGACC	LAEAGFDAFVEVSPHPVLT	GTCGGCATCGAGGCCACGCTGGACTCCGCGCTCCCGGCCGACGCCGGCGCGCTGCGTG ++++++	GIEATLDSALPADAGACVV	GGCACCCTGCGGCGGGACCGGCGGCCTGGCCGACTTCCACACCGCGCTCGGCGAGGCG	TLRRDRGGLA.DFHTALGEA	TACGCGCAGGGCGTGGAGGTCGACTGGAGCCCCGCCTTCGCCGACGCGCGGCCGGTCGAG	A Q G V E V D W S P A F A D A R P V E	CTGCCCGTCTACCCGTTCCAGCGGCAGCGGTACTGGCTGCCCATCCCCACCGGCGGCGC
Ω.	GA 25723	Ω	CA 25783	Ø	GT 25843	>	GC 25903	O	TA 25963	X	C1 26023
				SUE	STITUT	TE SI	HEET				

					70)/89				(FIG.2-69
I	26142		26202	ı	26262	1	26322	1	26382	i	264
V Y P F Q R Q R Y W L P I G G R	GCACGGGACGAGGACGACTGGCGCTACCAGGTCGTATGGCGGGAAGCCGAGTGGGAG ++++++) E D D D W R Y Q V V W R E A E W E	AGCGCTTCGCTGGCCGGACGCGTGCTGCTGGTGACCGGACCGGGCGTGCCGTCCGAGTTG	SLAGRVLLVTGPGVPSEL	TCGGACGCCATCCGAAGTGGACTGGAGCAGAGCGGTGCGACGGTCCTGACCTGCGACGTG	AIRSGLEQSGATVLTCDV	GAATCCCGTTCGACCATCGGCACCGCACTGGAGGCCGCCGACACCGACGCTCTGTCCACT	RSTIGTALEAADTDALST	GTGGTGTCGCTGTCCCCGCGACGCCGACGCCGTCGATCCGTCGCTGGACGCGCTCGCC	SLLSRDGEAVDPSLDALA	CTGGTCCAGGCCCTCGGAGCGGCCGGGGTCGAAGCACCGCTGTGGGTGCTGACCCGCAAC
I P	GCACGG	A R D	AGCGCT	8	TCGGAC	S	GAATCC	ы S	GTGGTG	> >	CIGGIC
	26083		26143		26203		26263		26323		26383
								SL	JBSTIT	ruti	SHE

ATCGGCGCGCACGTCGCCCGCTGGCCCGCTCGGGCGCCCGAGCACCTGGTGCTG

					71,	/89		! !	FIG.2-70	
ſ	26502	1	26562	ı	26622	1	26682	1 1	267	i
LVQALGAAGVEAPLWVLTRN -	GCCGTGCAGGTGGCCGACGGCGAACTGGTCGATCCGGCGCGCAGGCCATGGTGGGGCGGTCTC	AVQVADGELVDPAQAMVGGL.	GGCCGCGTGGTCGGCATCGAGCCGGGGGGGCGCTGGGGCGGTCTGGTGGACCTGGTCGAC (GRVVGIEQPGRWGGLVDLVD	GCCGATGCCGCGTCGATCCGGTCGCTGGCCGCGGTGCTGGCGGACCCCGCGCGGCGAGGAG	ADAASIRSLAAVLADPRGEE	CAGGTCGCGATCCGGGCGGACGGGATCAAGGTGGCGAGGCTCGTGCCCGCCC	Q V A I R A D G I K V A R L V P A P A R	GCCGCACGCACCCGCTGGAGCCCTCGCGGCACCGTGCTGGTCACCGGCGGCGCACCGGAGGG	A A R T R W S P R G T V L V T G G T G G
	26443		26503		26563		26623		26683	
			01	IDOT	T 172 (T			

SUBSTITUTE SHEET

-+ 26802		GGCAGGCGCGGTGCCGACGCACCCGGCGCGTCCGAGCTGAGGGAGG	ı	CCGCCTGCGACGTCGCCGACCGGGCGCGCTCGAAGCGGTG	· >	CTCGCCGCGGAGCGCCCGAGGGACGCACGGTCAGCGCCGTGATGCACGCGGGGGGTT	· ^	TCCACGTCCACGCCCCTCGACGTCACCGAAGCCGAGTTCACCGAGATCGCCGACGTG ++) >	AAGGTGCGCGCCCCGTCAACCTGGACGAGCTCTGCCCGGACCTCGACGCGTTCGTGTTG F1G.2-71	١
1	н	CCGC	4	AAGC	Æ)990	U	500	Ω	TCG	
į	>	TGA(H	TCG	ធ	990	4	TCG	4	CGI	<u>[</u>
+	H	AGCT -+	H	-+- 	H	CACGC	∢	SAGA	ii ii	GACGCG +	6
i	#	AGG	臼	090	K K	TGC	=	VCC0	H	CHC	,
	ធ	999	ជា	999	&	TGA	Σ	rrcz	E4	SACC	í
+-+	A	TGA +	α,	;ACC	<u>د</u> د	+ 52252	A	3AG1	ធ	+	•
i i	Ů	GAGCTG	<u>н</u>	500	Ω	000	S	3000	A H	CTGC(•
 	S	500	回	TCG	4	STCA	>	SAAG	臼	CHC.	,
 	M M	3CG1	S	SACG	, D	CACGG	H	ACCC	EH	CGAGCT	ſ
1.1	L A	25256	9	16CC	່ ບ	7. J	α.	CIC	i.	GAC	
	W	3000	<u>۵</u>	1 2 2 3 3 1	Æ	3GA(<u>ප</u>	SAC	Д	CTG	
ļ. 	R P	SCAC	A H	3000	4	GAG	ы	GAC	<u> </u>	AAC	
†	Æ	3AC(+	А	ATCG	. Н	+ - 2009	Æ	CTCG	, H	GTCA	;
Í	>	000	A	ACC	H	292	ø	0	Ω4.	ACC	E
İ	#	GGT	U	GTG		GAG		ACG	H	299;	(
+-	Æ	- + -	æ	GGCACGGCGTGACCATCG	. ტ	+- .60067	Ø	TCC:-+-	တ	-+ 3035	(
1	U	AGG	æ	ACG	H	၁၁၅	Ø	ACG	EH	GTG	;
1	н	299	ŋ	299		CIC	н	HCC 	လ	AAG	;
26743		26803		とりなりつ		26923		26983		27043	

					FIG.2-7
27162	27222	27282	27342	27402	27
TTCTCCTCCAACGCGGCGTGTGGGGCAGTCCGGGGCTCGCCTCCTACGCGGCGCCAAC 3+++++++	GCCTTCCTCGACGCTTCGCGCGCGCGCGCGCGAGGGCGCGCCGCTGACGTCCATC 3+++++++-	GCCTGGGGGCTCTGGGCCGGGCAGACATGGCCGGGGACGAGGGCGGCGAGTACCTGCGC 3+++++++	AGCCAGGGCCTGCGGGCCATGGACCCGGATCGGGCCGTCGAGGAACTGCACATCACCCTC 3+++++++	GACCACGGTCAGACGTCGTGGTCGTGGACATGGATCGCAGGCGGTTCGTCGAGCTG 13++++++++	TTCACCGCGCCCGGCCGCCGCTGTTCGACGAGATCGCCCGGTGCCCGGGCGGAAGCC 13+++++++
27103	27163	27223	27283	27343	27403
				SUBSTITU	JTE SHEET

•					7	74/89			1	FIG.2-73	
27522	ı	27582	i	27642	1	27702	1	27762	1	278	ı
CGGCAGAGCGAGGAGGGCCCGGCGCTCGCCCAGCGGCTCGCGGCGCTGTCGACGGCCGAG	ROSEEGPALAORLAALSTAE	AGGCGCGAGCACCTCGCCCACCTGATCCGCGCCGAGGTCGCCGCGGTGCTCGGCCACGGC	RREHLAHLIRAEVAAVLGHG	GACGACGCGGCGATCGACCGCGACCGCCTTCCGCGACCTCGGCTTCGACTCCATGACC	D D A A I D R D R A F R D L G F D S M T	GCCGTCGACCTGCGGAACCGGCTCGCCGCGGTGACCGGGGTGCGGGAAGCCGCGACGGTG	AVDLRNRLAAVTGVREAATV	GICITCGACCACCGACCATCACCCGGCTCGCCGACCACTACCTGGAGCGGCTCGTCGGC	V F D H P T I T R L A D H Y L E R L V G	GCAGCAĠAGGCGGAGCAAGCCCCGGCGCTCGTGCGCGAGGTGCCGAAGGATGCCGACGAC 3+++++++	A A E A E Q A P A L V R E V P K D A D D
27463		27523	_	27583		27643	-	27703		27763	
			S	UBSTI	TUTE	SHEE	Γ .				

75/89										F 1 G. Z – 7.
27882	ı	27942	ì	28002	ı	28062	ı	28122	1	28
CCGATCGCGATCGTCGGCATGGCCTGCCGCTTCCCCGGCGGCGTGCACAACCCCGGTGAG	PIAIVGMACRFPGGVHNPGE	CTGTGGGAGTTCATCGTCGGCCGCGGAGACGCCGTGACGGAGATGCCCACCGACCG	LWEFIVGRGDAVTEMPTDRG	TGGGACCTCGACGCGCTGTTCGACCCCGACCCGCAGCGCCACGGAACCAGCTACTCGCGA	WDLDALFDPDPQRHGTSYSR	CACGGCGCGTTCCTCGACGGGCCGCCGACTTCGACGCGCGTTCTTCGGGATCTCGCCG	HGAFLDGAADFDAAFFGISP	CGCGAGGCGCTGGCGATGGACCCGCAGCAGCGCCAGGTCCTGGAAACGACGTGGGAGCTG	REALAMDPQQRQVLETTWEL	TICGAGAACGCCGGCAICGACCCGCACTCGCIGCGGGGCAGCGACACCGGCGICTICCIC
27823		27883		27943		28003		28063		28123
				ATIT! 17	E 61	IFFT				

•					76	6/89				1	1/ 5
1	28242	ı	28302	ı	28362	ı	28422	1	28482	i	FIG. 2-/5 28542
ENAGIDPHSLRGSD'TGVFL	GGCGCCGCGTACCAGGGCTACGGCCAGGACGCGGTGGTGCCCCGAGGACAGCGAGGGCTAC	A A Y Q G Y G Q D A V V P E D S E G Y	CTGCTCACCGGCAACTCCTCCGCCGTGGTGTCCGGCCGGGTCGCCTACGTGCTGGGGCTG	LIGNSSAVVSGRVAYVLGL	GAAGGCCCCGCGGTCACGGTGGACACGGCGTGTTCGTCGTCGTTGGTGGCCTTGCATTCG	GPAVTVDTACSSSLVALHS	GCGTGTGGGTCGTTGCGTGACGGTGACTGCGGTCTTGCGGTGGCCCGGTGGTGTGTCGTG	CGSLRDGDCGLAVAGGVSV	ATGGCGGGCCCGGAGGTGTTCACCGAGTTCTCCCGCCAGGGCGGCTTGGCCGTGGACGGG	AGPEVFTEFSRQGGLAVDG	CGCTGCAAGGCGTTCTCCGCGGAGGCCGACGGCTTCGGTTTCGCCGAGGGCGTCGCGGTG ++++++-
[E4	GG 28183	Ŋ	CT 28243	ı	<i>G</i> 2 28303	ш	28363	æ	AT 28423	Σ .	CG 28483

					77/89)		!	FIG.2-76	
								ļ	.2-	
	28602		28662		28722		28782		F1G 28842	
1	28	1	28	1		ı		1		ı
>	GTCCTGCTCCAGCGGTTGTCCGACGCCCGCAGGGCGGGTCGCCCAGGTGCTCGGCGTGGTGTC ++++++	>	GCGGGCTCGGCGATCAACCAGGACGGCGCGAGCAACGGTCTCGCGGCGCCCGAGCGGCGTC +++++++	>	GCCCAGCAGCGCGTGATCCGCAAGGCGTGGGCGCGTGCGGGGGATCACCGGCGCGCGGATGTG	>	GCCGTGGTGGAGGCGCATGGGACCGGTACGCGGCTGGGCGATCCGGTGGAGGCGTCGGCG	Æ	TTGCTGGCTACTTACGGCAAGTCGCGCGGGTCGTCGGGCCCGGTGCTGCTGGGTTCGGTG	>
æ)GT(>	366	ט	3GA	Ω	GTC	S	7.T.	ဟ
>	999	U	3AG(S	Ö	æ	360	A	999	ပ
G	010	H	200	а) 1992	Ŋ	3GA(ធ	GCT-+	H
ជា	AGGTGCTCGGC +	>	92292929	Æ	CACC	Ħ	3GT(>	GCTGCTGG	1
Æ	CAC	œ	2000	4	SAT(H	I C I	Д	3GT(>
Įті	GGGGGGTCGCC	~	AACGGTCTC	ы	3GG 	ტ	CGA'	Q	GTCGGGCCCGGT	Δ ₄ .
Ö	36GJ	ტ	+	G	IGC(æ	GGTACGCGGCTGGGCGA	ტ	+ 9999	<u>ග</u>
[z _i	3600	Æ	CAA	z	300	æ	GCT	н	GTC	လ
ပ	SAGO	ĸ	SGCGCGAGC	လ	CGTGGGCG	Ø	929	œ	GCGCGGGTC	လ
Ω	ACGCCCGCA	ĸ	: + - : + -	æ	GTG(-+-	3	TAC -+-	Ħ	-+-	Ŋ
Ø) 1 1 1 1	A	266	Ŋ) 1 1 1	æ	990	Ŋ	 909	K
Ħ	CGA(Q	GGA(Ω	CAA	×	GAC	E→	GTC	လ
K	TGTC(လ	CCA	ø	CCG++	K	TGG +	Ŋ	CAA +	×
S	GTT(H	TCAAC	z	GATC(н	GCAT	# ,	TACGGC	೮
Įω	GTCCTGCTCCAGCGGT	ď	GAT	H	GCCCAGCAGCGCGTG	>	GCCGTGGTGGAGGCG	K	TTA 	×
Æ	CCA	a	GCGGGCTCGGCGA	æ	 929	ĸ	GGA 	ជា	TGCTGGCTACT	E
×	GCT(႕	CHC +	ဟ	GCA +	ø	GGT 	>	+ 299	A
υ	CCT	н	366	ტ	CCA	Ø	CGT	>	GCT	ı
A	GT(>	Ü	8	Ü !	æ		Æ		н
	28543		28603		28663		28723		28783	

28843	AAGTCGAACATCGGTCACGCGCGGCGGCGGGGGGTGTCGCGGGGGGTGTCAAGGTGGTGGTC 43+++++++		28902
	K S N I G H A Q A A A	G V A G V I K V V	1
28903	CTGGGGTTGAACCGCGG	CCTGGTGCCGCCGATGCTCTGCCGCGCGAGCGGTCGCCGCTG	28962
	LGLNRGLVPPM	LCRGERSPL	ı
28963	ATCGAATGGTCCTCGGG	TGGTGTGGAACTTGCCGAGGCCGTGAGCCCGTGGCCTCCGGCC	29022
	IEWSSGGVELA	EAVSPWPPA	1
29023	GCGGACGGGGTGCGCCGGGCCGGTC		29082
	ADGVRRAGVSA	FGVSGTNAH	i
29083	GTGATCATCGCCGAGCCCCCGGAG		29142
	VIIAEPPEPE	PLPEPGPVGV	ı
•	CIGGCCGCIGCGAACIC	CTGTCGGCCAGGACCGAGACCGCGTTGGCA	FIG.2-7
29143			70767

					7	9/89					FIG.2-78
ı	29262	ì	29322	1	29382	ı	29442	ı	29502	ı	FIG. 29562
LAAANSVPVLLSARTETALA	GCGCAGGCGCGCTCCTGGAGTCCGCAGTGGACGACTCGGTTCCGTTGACGGCATTGGCT	A Q A R L L E S A V D D S V P L T A L A	TCCGCGCTGGCCACCGGACGCCCCACCTGCCGCGTCGTGCGGCGTTGCTGGCAGGCGAC	SALATGRAHLPRRAALLAGD	CACGAACAGCTCCGCGGGCAGTTGCGAGCGGTCGCCGAGGGCGTTGCGGCTCCCGGTGCC	HEQLRAVAEGVAAPGA	ACCACCGGAACCGCCTCCGCCGGCGCGTGGTTTTCGTCTTCCCAGGTCAGGGTGCTCAG ++++++	TTGTASAGGVVFVFPGQGAQ	TGGGAGGGCATGGCCCGGGGCTTGCTCTCGGTCCCCGTCTTCGCCGAGTCGATCGCCGAG	WEGMARGLLSVPVFAESIAE	TGCGATGCGGTGTTGTCGGAGGTGGCCGGGTTCTCGGCCTCCGAAGTGCTGGAGCAGCGT +++++++
	29203		29263		29323		29383		29443	•	29503

						80/8	9			FIG 2-70	
1	2962	1	29682	ı	29742	1	29802	ı	29862	י בו	2
CDAVLSEVAGFSASEVLEQR -	CCGGACGCCCGTCGCTGGAGCGGGTCGACGTCGTACAGCCGGTGTTGTTCTCCGTGATG	PDAPSLERVDVVQPVLFSVM -	GTGTCGCTGGCGCGGCTGTGGGGCGCTTGCGGAGTCAGCCCCTCGGCCGTCATCGGCCAT	V S L A R L W G A C G V S P S A V I G H	TCGCAGGGCGAGATCGCCGCCGCGGTGGTGGCCGGGGTGTTGTCGCTGGAGGACGGCGTG	SQGEIAAAVVAGVLSLEDGV	CGCGTCGTGGCCCTGCGCGCGAAGGCGTTGCGTGCGCTGGCGGGCAAGGGCGGCGATGGTC	RVVALRALRALAGKGGMV	TCGTTGGCGGCTCCCGGTGAACGCGCCCGCGCGCTGATCGCACCGTGGGAGGACCGGATC	SLAAPGERARALIAPWEDRI	TCCGTCGCGGCGGTCAACTCCCCGTCCTCGGTCGTGGTCTCCGGCGATCCGGAGGCGCTG
	29563		29623		29683		29743		29803		29863

CACCCGGTGCTCACCGCGGCGGTGCAGGAGATCGCCGCGGGACGCCGTGGCCATCGGGTCG

						81 /	/89	(-80	
1	29982	ſ	30042	ı	30102	ı	30162	1	FIG.2-80	1
V A A V N S P S S V V V S G D P E A L	GCCGAACTCGTCGCACGTGCGAGGGCGTGCGCGCCCAAGACGCTCCCGGTGGAC	ELVARCEDEGVRAKTLPVD	TACGCCTCGCACTCCCGCCACGTCGAGGAGATCCGCGAGACGATCCTCGCCGACCTCGAC	ASHSRHVEEIRETILADLD	GGCATCTCCGCGCGGCGTGCCGCCATCCCGCTCTACTCCACGCTGCACGGCGAACGGCGC	I S A R R A A I P L Y S T L H G E R R	GACGGCGCCGACATGGGTCCGCGGTACTGGTACGACAACCTGCGCTCCCAGGTGCGCTTC	G A D M G P R Y W Y D N L R S Q V R F	GACGAGGCGGTCTCGGCCGCCGTCGCCGACGGTCACGCCACCTTCGTCGAGATGAGCCCG	E A V S A A V A D G H A T F V E M S P
ν	29923	R	TA 29983	X	30043	ຶ	G 30103	Q	G. 30163 -	Q

						82/	8 9		FIG. 2-8			
30282	ı	30342	ı	30402	ı	30462	ı	30522	į	F16.	i	
+ -	ω·	CGGAGGAGCACCTGATCGCCGAGCTCGCCCGGGCGCACGTGCAC	Ħ	GGCGGAACGTCTTCCCGGCGCACCTCCGGTGGCGCTGCCCAAC	z	AGCGGTACTGGCTCGCGCGGAGGTGTCCGACCAGCTCGCCGAC	Ω	AGCCGCTACCGCGTCGACTGGCGACCGCTGGCCACCACGCCGGTGGACCTGGAAGGCGGC	ڻ و	CCGCACCGGAGTCGCTGACCAGCGCAGTCGAGAAGGCCGGAGGC	ტ	
i 1	U	3TG	>	20	Ωı	300	æ) 1 1 1	v	GGA	Ŋ	
i	Н	CACC	H	CTG	H	CHC	H	SAA	ធ	Ο I Ο I	A	
		900		900	ر لا	SAG(α	CTG	1	CGAGAAGGCCGG	×	
+	A	+	A.	GTGGC +		ACC +-	٥	3ACC		3AG2	ы	
1	>		K.	 590	>	000		TGG	Δ	TCC		
1	K	JCG 1	K	CTC	Д	TGT 	တ	990	>	CAG	>	
ļ	Ω	1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -	1	CAC	Сı	AGG +	>	+ +	Д	- - +	A	
Ī	Ø	+	ப	GCGGCAC	K	7992	ជា	CCA	H	CCA	S	
į	æ	95	K	999	K	Σ̈́̈́ς	Д	CA	H	rga(H	
	н	GAT	H	CTTCCCG -+	Δ	990	Æ	999	K	, GC	1	
+	ជា	ACCT	1	CTT	لعر	GCT -+-	Н	+ +	H	AGTCGCTGACCAGCGC	လ	
	a	GCA	Ħ	CGT	>	CTG	3	ACC	Ф	GGA	ជ	
	>	3GA	ធា	SAA	z	GTA	×	909	¤	ACC 	Δı	
į	K	CGGA(-+	ធ	+- 92999	×	AGCG	ĸ	ACTG	3	28C.	A	
	Æ	900	æ	7.TG(3	CA	α	CGA	Ω	STC	လ	
į	H	ACC	H	GAC	Ω	ည	Д)GT(>	999	ტ	
	H	GAC	Ω	GTG	>	GAG	ធ	990	æ	CAC	Ħ	
+-	>	950	æ	-+	Ø	TTC -+-	Ĺij	CTAC	≽ı	GTC -+-	>	
	Д	CAC	×	GTG	>	900	Д	292 292	~	CTG	ы	
	=	CTGCACCGCGACACCG	H	GGCGTGGCCGTGGACT	ິບ	TACCCGTTCGAGCCCC	₩	AGC	S	TTCCTGGTCCACGGGT	ĹĽI	
30223	-	30283	•	30343	•	30403	•	30463	-	30523		

•.						83/8	19		C	70 -	
30642	1	30702	ı	30762	1	30822	ı	30882	, ,	FIG.C - 0C 30942	i
CGCGTCGTGCCGGTCGCCTCGGCCGACCGCGAAGCCTCGGCGGCCCTGCGGGAGGTGCCG	ASADREASAALREVP	GGCGAGGTCGCCGGCGTGCTCTCGGTCCACACCGGCGCCGCAACGCACCTCGCCCTGCAC	V L S V H T G A A T H L A L H	CAGTCGCTGGGTGAGGCCGGCGTGCGGGCCCCGCTCTGGCTGG	A G V R A P L W L V T S R A V	GCGCTCGGGGAGTCCGAGCCGGTCGATCCCGAGCAGGCGATGGTGTGGGGGTCTCGGGCGC	E P V D P E Q A M V W G L G R	GACCCCGGAACGGTGGGCGGTCTGGTGGACCTGCCCGCCGAACCC	TPERWGGLVDLPAEP	GCGCCGGGGGACGGCGAGGCGTTCGTCGCCTCGGCGCGCGGGGGCGGCCACGAGGACCAG	EAFVACLGADGHEDQ
TGCCGG7	>	GGTCGCCG(A G	GCTGGGTG1	ម ២ -	CGGGGAGT(<u>ធ</u>	GGGCCTGGA	i E	GGGGGACG	D C
CGCGTCG	В V V	GGCGAGG	G E V	CAGTCGC	S G	GCGCTCG	A L G	GTCATGG	Σ Σ	9922929	À P G
30583	-	30643	•	30703		30763		30823		30883	

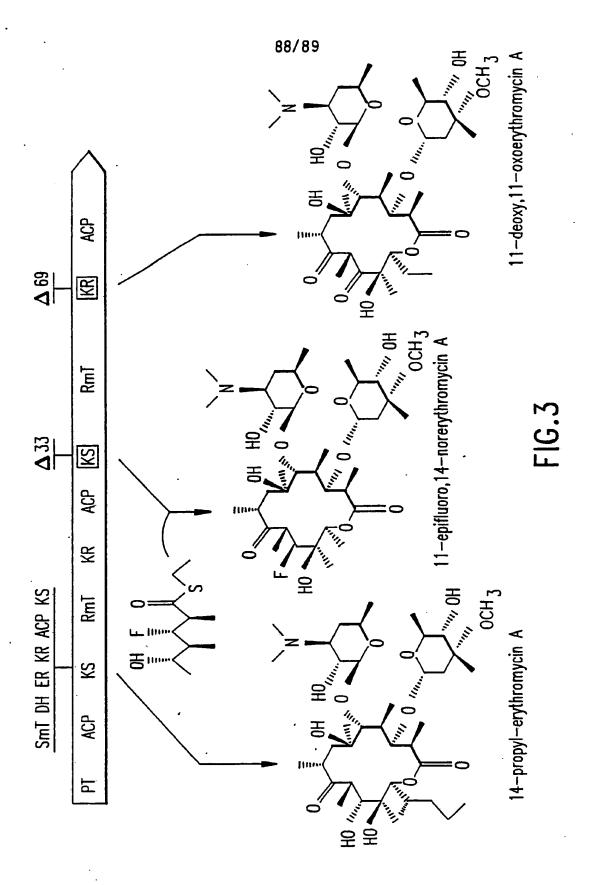
			8	4/89			FIG 9_83	3
31002	31062	31122	1	31182	ŧ	31242	, 514 -	ω.
GTCGCGATCCGTGACCACGCCCGCTACGGCCGCCGCCTCGTCCGCCCCCCCC	CGCGAGTCGAGCTGGGAGCCGGCGGGCACGGCG	GGCGGCCACGTCGCCACCTCGCCAGGTGCGGGGTGGAGGACCTGGTGCTGGTC	G G H V A R H L A R C G V E D L V L V S	AGGCGCGCGCGTCGACGCTCCCGGCGGCCGAGCTGGAAGCCGAACTGGTCGCCCTCGGC	RRGVDAPGAAELEAELVALG	GCGAAGACGACCATCACCGCCTGCGACGTGGCCGACCGCGAGCAGCTCTCCAAGCTGCTG	AKTTITACDVADREQLSKLL	GAAGAACTGCGCGGGCAGGGACGTCCGGTGCGGACCGTGCACACCGCCGGGGTGCCC
30943	31003	31063		31123		31183		31243

PCT/US92/00427

	·					85 / 8	9				FIO.Z-84 s62
3	31363	7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	31422	1	31482	1	31542	1	31602	l L	710.
	GAATCGAGGCCGCTGCACAGATCGGCGAGCTGGAGTCGGTCTGCGCGCGC	ESRPLHEIGELESVCAAKV	GGGGCCCGGCTGCTCGACGAGCTGTGCCCGGACGCCGAGACCTTCGTCCTGTTCTCGTCC 53	GARLLDELCPDAETFVLFSS	GGAGCGGGGGTGTGGGGCAGTGCGAACCTCGGCGCCTACTCCGCGGCCAACGCCTACCTC	G A G V W G S A N L G A Y S A A N A Y L	GACGCGCTGGCCCACCGCCGCGTGCGGAAGGCCGTGCGGCGACGTCCGTC	DALAHRRAEGRAATSVAWG	GCCTGGGCGGGCGAGGGCATGGCCACCGGCGACCTCGAGGGGGCTCACCCGGCGCGCCTG	AWAGEGMATGDLEGLTRRGL	CGCCCGATGGCGCCCCGAGCGCGCGATCCGCGCGCTGCACCAGGCGCTGGACAACGGCGAC
	31303	,) i	31363		31423		31483		31543		31603

						86/8	9			0	F10, Z - 80
ı	31722	1	31782	t	31842	ı	31902	1	31962	<u>C</u> L	F10.
PMAPERAIRALHQALDNGD	ACGIGCGITICGAICGCCGACGICGACIGGGAGGCCIICGCGGICGGCIICACCGCCGCC	CVSIADVDWEAFAVGFTAA	CGGCCGCGTCCGCTGCACGAGCTCGTCACGCCGGCGGTGGGGGGCCGTCCCCCGCGGTG	PRPLLDELVTPAVGAVPAV	CAGGCGGCCCCGGCGCGGGAGATGACGTCGCAGGAGTTGCTGGAGTTCACGCACTCGCAC	AAPAREMTSQELLEFTHSH	GTCGCGGCGATCCTCGGGCATTCCAGCCCGGACGCGGTCGGGCAGGACCAGCCGTTCACC	таатысназерау 6 орорет	GAGCTCGGCTTCGACTCGCTGACCGCGGTCGGGCTGCGCAACCAGCTCCAGCAGGCCACC	E L G F D S L T A V G L R N Q L Q Q A T	GGGCTCGCGCTGCCCGCGACCCTGGTGTTCGAGCACCCCCACGGTCCGCAGGTTGGCCGAC
er.	A 31663 -	E	31723 -	X	31783	a	31843 -	>	31903	•	31963

					87,	'8 9			FIG.2-86	
	32082		32142		32202		32262		FI(32322	•
ı		1		1		i		1		ŀ
Q	acagcgggactcccgcccgggaagcgagcagcgctcttcgc ++++++	K	GCGTGTCGGGCAGGGTCCGGTCCTACCTCGACCTGCTGGCG	æ	AGCACTTCGACGGCTCCGACGGGTTCTCCCTCGATCTCGTG	>	GACATGGCCGACGGTCCCGGAGAGGTCACGGTGATCTGCTGCGCGGGAACGGCGGCGATC	H	CCCGGCTCGCCGGGGCGCTGCGCGGAATCGCTCCGGTTCGG	æ
Æ	FCF	H,	GCT	ы	TCT	, ,, ,	2992	Ø	1997	>
H	292	Æ	CCT	ы	CGA	Ω	0993	Ø	CTCC	ים
ĸ	CAG +	လ	+	Ω) +	h	3AAC -+	H	rcgc +	Æ
ಜ	GAG	S	CCT	H	CTC	ဟ)993	Ŋ	SAAS	н
>	AGC	æ	CTA	×	GTI	ĹĿ	3000	A	3000	Ŋ
H	GGA	ធ	GTC	လ	AGCACTTCGACGGCTCCGACGG	ប	SCTC	U	76CC	8
Д	-+	K	900	ĸ	100 Y	Ω	CTC	U	36CJ	1
Ξ	090	4	GGT	>	CIC	ဟ	GAJ	н	3660	K
ы	TCC	۵.	CAG	8	ອອວາ	O	3661	> .	9900	r
Ŀı	GGGACT	H	- + -	ပ	£50 + + -	Ω	CAC	H	555	Ø
>	992	Ŋ	GIC	S	CTT	Ĺų	GGT	>	GCT	7
1	CAG	S	CGT	>	GCA	Ħ	GAGA	ធ	999	ሺ
H	CTCGA	Ω	+ 999	Ö	CGA +	ធា	+ +	ტ	TCAC	H
4	GCT	H	255	Ø	900	~	TCC 	а	GTT	ក្រ
<u>α</u>	GCA	ø	GCA	ø	CII	Ĺ	9901	Ŋ	CGA	ចា
7	CACATAGGACAGCAGCTCG	a	GACGGCTACCGGCAGGCGG	~	GGGCTGTCGGACTTCCGCG	Ω	TGGCCGACGGTCCCG	Ω	TCCGGTCCGCACGAGTTCA	Ξ
Ø	AGGA	ტ	CIA	×	G7C 	S	660	A	11CC	Ω
H	CAT	н	 	ტ	GCT	1	CAT	Σ	990	O
Ū	CAC	Ħ		Ω		ပ		Ω		ທີ
	32023		32083		32143		32203		32263	



SUBSTITUTE SHEET

NUMBER	SITE	DISTANCE (Kb)a
1. 1	BamHl	-3.60
2	Pvull	-3.50
3	Pvull	-3.40
1	Pstl	-3.05
4 5	BamHl	-2.95
6	Xhol	-2.80
7	Pstl	-2.00
8	Hindll	-1.60
9	Sphl	-1.55
10	EcoRI	-1.50
11	Kpnl	-1.35
12	EcoRI	- 1.05
13	Smal b	-0.90
14	Sphl	-0.75
15	Kpnl	-0.65
16	Smal	-0.20
1		

FIG.4

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US92/00427

1. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)3 According to International Patent Classification (IPC) or to both National Classification and IPC IPC (5): A01N 43/22; A61K 31/71; C07H 17/08; C12N 1/21; C12P 19/62 US CL : 435/76, 252.35, 886; 514/29; 536/7.2 II. FIELDS SEARCHED Minimum Documentation Searched 4 Classification Symbols Classification System 435/76, 183, 252.35, 886; 514/29; 536/7.2 U.S. Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched 6 APS, BIOSIS, MEDLINE, BIOTECH. ABSTRACTS III. DOCUMENTS CONSIDERED TO BE RELEVANT 14 Citation of Document, 16 with indication, where appropriate, of the relevant passages 17 Relevant to Claim No. 18 Category* J. of Bacteriology, Volume 164, No. 1, issued October 1985, J.M. Weber et al, "Genetic Analysis of 1985, J.M. Weber et al, "Genetic Analysis of Erythromycin Production in Streptomyces erythreus", pages 425-433, See the entire document. 1-30 J. of Bacteriology, Volume 172, No. 5, issued May 1990, J.M. Weber et al, "Organization of a Cluster of Erythromycin Genes in Saccharomyces erythraea", pages 2372-2383, See the entire document. US, A, 4,874,748 (Katz et al) 17 October 1989, see 1-30 column 1, lines 41-59; column 3, lines 47-60. US, A, 4,935,340 (Baltz et al) 19 June 1990, see entire document. later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention Special categories of cited documents: 15 document defining the general state of the art which is not considered to be of particular relevance earlier document but published on or after the international filing date document of particular relevance; the claimed invention cannot be considered novel or cannot be document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) considered to involve an inventive step document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with document referring to an oral disclosure, use, exhibition or other means one or more other such documents, such combination being obvious to a person skilled in the art document published prior to the international filing date but later than the priority date claimed document member of the same patent family IV. CERTIFICATION Date of Mailing of this International Search Report² 30 MAR 1992 Date of the Actual Completion of the International Search² 20 MARCH 1992 Signature of Authorized Officer 20 Warnel International Searching Authority¹ ISA/US Dian Cook

Form PCT/ISA/210 (second sheet)(May 1986) &